

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
4 September 2003 (04.09.2003)

PCT

(10) International Publication Number  
**WO 03/072788 A1**

(51) International Patent Classification<sup>7</sup>: **C12N 15/63**,  
15/85, C07H 21/04

350 New Market Court, Chesterbrook, PA 19087 (US).  
**SCHULTZ, David C.** [US/US]; 5260 Hickory Drive,  
Lyndhurst, OH 44124 (US).

(21) International Application Number: PCT/US03/05347

(22) International Filing Date: 20 February 2003 (20.02.2003)

(74) Agents: **BAK, Mary E.** et al.; Howson & Howson, Spring  
House Corporate Center, P.O. Box 457, Sprinhouse, PA  
19477 (US).

(25) Filing Language: English

(81) Designated States (*national*): AU, CA, JP, US.

(26) Publication Language: English

(30) Priority Data:  
60/358,599 21 February 2002 (21.02.2002) US

(84) Designated States (*regional*): European patent (AT, BE,  
BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU,  
IE, IT, LU, MC, NL, PT, SE, SI, SK, TR).

(71) Applicant (*for all designated States except US*): **THE  
WISTAR INSTITUTE OF ANATOMY AND BIOL-  
OGY** [US/US]; 3601 Spruce Street, Philadelphia, PA  
19104 (US).

**Published:**

- *with international search report*
- *before the expiration of the time limit for amending the  
claims and to be republished in the event of receipt of  
amendments*

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **RAUSCHER,**  
**Frank J.** [US/US]; 360 Conestoga Road, Wayne, PA  
19087 (US). **AYYANATHAN, Kasirajan** [IN/US];

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

(54) Title: METHODS AND COMPOSITIONS FOR REVERSIBLY CONTROLLING EXPRESSION OF TARGET GENES IN CELLS

(57) Abstract: A method for producing a cell containing a stably silenced target gene involves introducing into a parent cell a nucleic acid molecule expressing a chimeric repressor fusion protein containing the Krüppel-Associated Box (KRAB) domain of the KOX1 protein, or a variant thereof that binds KAP1 and has DNA-dependant repressor activity, and a targeting sequence that binds to a selected target gene. The fusion protein also contains as a switch component, a ligand-dependent binding domain. This parent cell is cultured in the presence of an effective amount of the ligand for a sufficient time to repress expression of the target gene, and the parent cell is single cell cloned in the absence of ligand to obtain single subclones thereof. Clones containing stably silenced target genes show no expression of the target gene after greater than 40 population doublings. Such cloned cells, optionally containing reporter plasmids, are useful in research and drug screening.

WO 03/072788 A1



## METHODS AND COMPOSITIONS FOR REVERSIBLY CONTROLLING EXPRESSION OF TARGET GENES IN CELLS

### BACKGROUND OF THE INVENTION

5           This invention relates generally to the fields of genetic engineering and gene therapy. More specifically, the invention relates to methods and compositions that enable the manipulation of gene expression.

          A variety of engineered transcriptional repressors have been described by the inventors and others for use in suppressing gene expression in cells, such as by  
10   inhibiting gene-activated pathways relevant to disease. Inducible transcriptional repressor constructs that fuse desired repressors domains to DNA binding sequences are used to target a gene for suppression. Hormone binding domains from a variety of hormone binding receptors have also been employed in such fusion proteins to permit inducible activity of the repressor in the cell. Such a system is described in W. J.  
15   Fredericks *et al*, 2001 *Cancer Lett.*, 162:S23-S32; W. J. Fredericks *et al*, 2000 *Mol. Cell. Biol.*, 20(14): 5019-5013; and K. Ayyanathan *et al*, 2000 *Cancer Res.*, 60:5803-5814, among others. Generally, in the presence of the inducer or ligand that binds the hormone-binding domain, these constructs become activated, resulting in transient suppression of the targeted gene in the cell. Upon withdrawal of the ligand from the  
20   cell, the target gene regains its normal expression levels. In such systems, the ligand operates to turn expression of the target gene "on and off".

          Even in view of the existing gene manipulation systems and methods, there remains a continuing need in the art for additional methods and compositions useful in manipulating gene expression for diagnosis, gene therapy, drug screening and  
25   biological research.

### BRIEF DESCRIPTION OF THE DRAWINGS

          Fig. 1A is a schematic representation, and characterization of conditional transcriptional repressor proteins. Chimeric repressor proteins have in common, the  
30   PAX3 DNA binding domain (PAX3), i.e., the targeting sequence, composed of the Paired Box (PB) and Homeodomain (HD) motifs, and the tamoxifen mutant estrogen receptor hormone binding domain (ERHBD<sup>TM</sup>), i.e., the ligand binding domain. The

KRAB, KRAB(DV), SNAG, Engrailed, PLZF-BTB/POZ and WT1 repression domains were fused in frame to the N-terminus of the PAX3-DBD, thus generating the different RD-PAX3-HBD fusion proteins.

Fig. 1B is a bar graph illustrating the 4-OHT-dependent repression of a PAX3 responsive luciferase reporter gene by the RD-PAX3-HBD proteins in NIH3T3 cells of Example 1. Post-transfection, duplicate dishes of cells were treated with 0.1% ethanol (-OHT) or 500 nM 4-OHT (+OHT) for 24 hours. Cell lysates were assayed for luciferase and  $\beta$ -galactosidase activities. Fold repression was determined as the ratio of normalized luciferase activity of -OHT to normalized luciferase activity of +OHT treated cells.

Fig. 1C is a schematic demonstrating the specific association of the KPHBD protein with HP1 $\alpha$  and HP1 $\gamma$ , as described in Example 2. The KPHBD protein interacts with the KAP1 corepressor. The KAP1 corepressor in turn interacts with the chromoshadow (CSD) domain of the HP1 protein family. CD indicates chromodomain.

Fig. 2A is a graphical illustration of the strategy to generate cell lines that stably express an engineered repressor protein and luciferase from a chromatin-integrated reporter transgene. Chimeric repressor proteins (RD-PAX3-HBD) were constitutively expressed from a CMV promoter. This plasmid also contains a neomycin<sup>R</sup> gene, whose expression is driven by the SV40 immediate early promoter, for stable selection in mammalian cells. The reporter plasmid, CD19-TK-LUC-Zeo<sup>R</sup> contains six copies of a PAX3 DNA-binding motif, followed by a minimal HSV TK promoter, which controls basal expression of the luciferase reporter gene. This plasmid was engineered to contain a zeocin<sup>R</sup> gene expression cassette for stable selection in mammalian cells.

Fig. 2B are bar graphs illustrating the characterization of clonal cells containing a chimeric RD-PAX3-HBD repressor and a chromatin-integrated luciferase reporter gene. Luciferase activity was measured in clonal cells after 24-hour treatment with 0.1% ethanol (-OHT; black bars) or 500 nM 4-OHT (+OHT; dotted bars) and expressed as light units/O.D. of protein. The top graph represents the normalized luciferase activities of 5 independent "CL" clones. The second, third, and fourth graphs illustrate the normalized luciferase activities in independent clonal

populations of cells stably expressing the respective RD-PAX3-HBD protein and luciferase from a chromatinized reporter.

Fig. 2C is similar to Fig. 2B. The top graph illustrates the fold repression observed between the - and + 4-OHT treated cells for the same clones. The second, third, and fourth graphs illustrate the corresponding fold repression observed between -OHT and +OHT treated cells.

Fig. 2D is a bar graph illustrating the characterization of the KPHBD21 cell line. KPHBD21 cells were treated with growth medium containing either 0.1% ethanol (-OHT) or varying concentrations of 4-OHT (+OHT) for either 0 hour, 12 hours, or 24 hours, respectively. Cell lysates were assayed for luciferase activity and normalized to 1 O.D. protein. Fold repression represents the ratio of normalized luciferase activity of -OHT to normalized luciferase activity of +OHT treated cells.

Fig. 2E are bar graphs that illustrate 4-OHT dependent repression of luciferase gene transcription as determined by quantitative RT-PCR. Oligo-dT primed first strand cDNAs from -OHT (0.1% ethanol) or +OHT (500 nM 4-OHT) treated KPHBD21 cells were PCR amplified with specific primer-pairs for the luciferase, neomycin<sup>R</sup>, and zeocin<sup>R</sup> genes (see Fig. 2A). Aliquots of the PCR reactions were removed at the indicated cycles and electrophoresed in 1.5% agarose gels, Southern-blotted with probes specific for luciferase, neomycin<sup>R</sup>, and zeocin<sup>R</sup> genes, and autoradiographed for equal duration (gels not shown). The signals were quantified by phosphorimager analysis and the arbitrary units plotted. Filled squares, -OHT; Filled circles, +OHT.

Fig. 3 is a schematic of the plasmids and the transgenes amplified by the primer pairs of Table 1 in Example 6. Row #s associated with Table 1 are shown beneath the relative fragment positions.

Fig. 4A is a diagrammatic representation of the mouse NT2-KRAB zinc finger protein. The N-terminal leucine-rich SCAN domain extends from amino acids 56 to 121, while the adjacent KRAB box is represented by amino acids 217 to 257. The C-terminal nine zinc fingers extend from amino acids 380 to 678.

Fig. 4B shows the genomic structure of *RXRβ* and *Coll1a2* genes. The oligonucleotides used in chromatin immunoprecipitation-PCR (ChIP-PCR)

experiments, their relative locations, and the sizes of the amplified fragments are indicated. E1 to E37 represent the exons of *Col11a2* gene.

Fig. 5 is a schematic illustration of the KAP-1 corepressor that interacts with the putative histone methyltransferase SETDB1. The oligomerization and KRAB binding domain map to the RBCC region of KAP-1. The chromoshadow domain of the HP1 family of chromosomal proteins directly binds to a PxVxL SEQ ID NO: 45 motif in KAP-1. The PHD finger and bromodomain of KAP-1 form a cooperative repression domain that interacts with Mi-2 $\alpha$  and SETDB1.

Fig. 6A illustrates that the KPHBD protein induces stable repression of the luciferase transgene. Duplicate dishes of KPHBD21 cells were treated with either 0.1% ethanol (-OHT) or 500 nM 4-OHT (+OHT) continuously for 24 or 48 hours, respectively. One dish from each of the - and + OHT treated cells was harvested at the end of induction. The remaining dishes were extensively washed to remove residual inducing agent. Pairs of dishes (-/+ OHT) were harvested at the indicated time points post 4-OHT removal. Luciferase activities were measured, normalized to protein concentration, and fold repression was determined (Filled Bars, 24-hour 4-OHT treatment; Patterned Bars, 48-hour 4-OHT treatment).

Fig. 6B illustrates the results produced when duplicate dishes of KPHBD21 cells were treated with either 0.1% ethanol (-OHT) or 500 nM 4-OHT (+OHT) continuously for either 2, 4 or 6 days. One dish from each - and +4-OHT treated cells was harvested at the end of induction and the remaining dishes were washed as above. At the end of 4th day of washings, the near confluent cells were trypsinized, re-plated in fresh dishes and washed for a maximum of 4 additional days. Corresponding - and +OHT treated cells were harvested daily and normalized luciferase activities were determined. Fold repression represents the ratio of normalized luciferase activity determined for -OHT versus normalized luciferase activity determined for +OHT treated KPHBD21 cells (Filled diamonds, 2 days of 4-OHT treatment; Filled circles, 4 days of 4-OHT treatment; Filled squares, 6 days of 4-OHT treatment).

Fig. 6C illustrates the treatment of the K(DV)PHBD3 (Open circles) and K(DV)PHBD22 (Closed circles) clonal cell lines with 500 nM 4-OHT continuously for 4 days, followed by removal of inducing agent and extensive washings as described in Figs. 6A and 6B.

Fig. 6D illustrates the treatment of the SPHBD11 (Closed circles) and SPHBD20 (Open circles) clonal cell lines with 500 nM 4-OHT continuously for 4 days, followed by removal of inducing agent and extensive washing as described in Figs. 6A and 6B.

Fig. 7A is a sectorial analysis of luciferase gene expression in KPHBD21 cells. An experimental scheme was designed to measure sectorial expression of the luciferase gene in KPHBD21 cells. The parental KPHBD21 cell line was sub-cloned by limiting dilution. Three subclones (KPHBD21-8, KPHBD21-39, KPHBD21-49) that possess stable levels of luciferase activity were treated with either 0.1% ethanol (-OHT) or 500 nM 4-OHT (+OHT) continuously for 4 days. Single cell sub-clones were generated by limiting dilution, propagated for ~40 population doublings, and assayed for normalized luciferase activity.

Fig. 7B shows the normalized luciferase activities for single cell progeny derived from KPHBD21-8, -39, or -49, respectively, as used in Fig. 7A. Each open square represents an independent single cell subclone of the parental line treated with 0.1% ethanol (-OHT). Each open diamond represents an independent single cell subclone of the parental line treated with 500 nM 4-OHT (+OHT). The total number of single cell subclones generated and analyzed under this experimental condition is indicated at the bottom for each parental line.

Fig. 7C is a bar graph showing the results of the clones 39-40 and clones 39-45 tested for their basal luciferase activities at approximately two-month intervals (the differently marked bars indicate the time periods).

Fig. 8A is a bar graph illustrating the basal luciferase activities of the active clone 39-45 and the silent clone 39-40 prior to ChIP experiments.

Fig. 8B graphically depicts the VPDBD plasmid containing VP16 acidic activation domain (amino acids 1-98) fused in frame with the PAX3 DNA binding domain (amino acids 99-479). The graph shows the luciferase activities and normalized activities with  $\beta$ -galactosidase light units of cell lysates from the silent clone transiently transfected with indicated concentrations of the VPDBD plasmid.

Fig. 8C is a bar graph demonstrating the results of the treatment of the silent (clone #39-40) and active (clone #39-45) clones with the 5-azacytidine (5AZA) and trichostatin-A (TSA) either alone or in combination in a sequential order for indicated

durations. Lysates were assayed for luciferase activities and the values were normalized to protein concentration.

Fig. 8D illustrates a representative selection of recombinant clones from sodium bisulfite-genomic sequencing of clone #39-40 and clone #39-45, that were PCR amplified using UMS1 (sense) and UMA1 (antisense) primers present in regions free of any CpG residues (to avoid any preferential amplification). Hanging diamonds represent the positions of CpG residues. Open squares represent the unmethylated CpG. Filled squares denote the methylated CpG residues. Upward arrows mark the CpG residues that are preferentially methylated in the genomic clones of clone #39-40.

Fig. 9 is a schematic illustration of the SETDB1 protein, a histone H3 specific methyltransferase. The position of the pre-SET, SET, and post-SET ('C') homologies at the COOH-terminus are indicated. The 347 amino acid insertion in the SET domain is indicated by the gray box. MBD represents a CpG DNA methyl binding domain. The relative position of single amino acid substitutions synthetically introduced into the MBD, pre-SET, SET, and post-SET domains is indicated. The minimal KAP-1 interaction domain (KID) is defined by amino acids of SETDB1 present in two-hybrid clone KIP21. The region of SETDB1 (amino acids 1 to 377) used to raise polyclonal and monoclonal antibodies is illustrated.

Fig. 10A are four schematic diagrams of HP1 $\alpha$  showing the domain organization (CD-chromodomain; CSD-chromoshadow domain) of this protein family and the relative position of the V21M and I165K mutations (Lechner et al. 2000 Mol. Cell. Biol., 20:6449-6465).

Fig. 10B is a bar graph illustrating the effect of histone modifications on the enzymatic activity of SETDB1. One microgram of unmodified or acetylated (K9-Ac, K14-Ac, K9, K14-Ac), phosphorylated (S10-phos) or methylated (K4-diMe, K9-diMe) peptides corresponding to NH2-terminal tail of histones H3 or H4 were used as substrates in the *in vitro* methylation assay with FLAG-purified SETDB1. Methylation was quantified via a filter binding assay and represented as raw C.P.M. incorporated.

Fig. 11 illustrates that KRAB: KAP-1 repression system targets SETDB1 and enhances H3-K9 methylation and HP1 recruitment to promoters of transcriptionally

silenced genes. A schematic representation shows a two plasmid system used to create a stably integrated luciferase transgene in NIH/3T3 cells that is regulated by a heterologous KRAB repressor protein. Numbered arrow sets represent the relative position of PCR primers used for PCR amplification of DNA retained by ChIP.

5

## SUMMARY OF THE INVENTION

In one aspect, the invention provides a method for producing a cloned cell containing a stably silenced target gene. This method involves introducing into a parent cell a nucleic acid molecule expressing a chimeric repressor fusion protein containing the Krüppel-Associated Box (KRAB) domain or a variant thereof that binds KAP1 and has DNA-dependent repressor activity, and a targeting sequence that binds to a selected target gene. The fusion protein also contains a component that acts as a "switch" to turn off or on the binding of the targeting sequence. In one embodiment, this switching component is a ligand-dependent binding domain, such as that of a nuclear hormone receptor. This domain controls the activity of the targeting sequence, making the binding inactive until a ligand is introduced into the system. The fusion protein is under the control of regulatory sequences capable of directing expression thereof in the parent cell. This parent cell is cultured in the presence of an effective amount of the ligand (inducer) for a sufficient time to permit binding of the targeting sequence and initiate thereafter a stable repression of expression of the target gene mediated by the KRAB domain. This repression can exist even in the absence of inducer. In one embodiment, the parent cell is single cell cloned in the absence of ligand to obtain single subclones thereof. About 30 percent of such clones do not express the target gene even after the cell clone has performed greater than about 50 cell population doublings in the absence of the ligand.

In another aspect, the invention provides a cloned cell containing a stably silenced target gene. This cell is desirably produced by the above method.

In a further aspect, the invention provides a screening method for identifying a test molecule that activates the expression of a stably silenced target reporter gene. A cloned cell of this invention or progeny thereof is contacted with a test molecule. The cell or progeny thereof are then monitored for expression of a target reporter gene.



Expression of the target gene in the presence of the test molecule indicates that the test molecule is able to reactivate expression of the silenced reporter gene.

In yet a further aspect, the invention provides a method for manipulating expression of a target gene in a cell by introducing to the cell containing said target gene an effective amount of a composition comprising a nucleic acid molecule expressing the above-described chimeric repressor fusion protein. Expression of the target gene is thereafter stably silenced by transiently contacting the cell with an effective amount of the ligand for a sufficient time to stably suppress expression of the target gene in subsequent progeny of the transfected cell in the absence of ligand.

In another embodiment of this method, expression of the silenced gene may be turned back "on" in progeny cells by contacting the cells with a molecule that disrupts the binding at the target gene of a complex mediated KAP1, HP1 and SETDB1.

In another aspect, the invention provides a method of producing a knock-out, non-human animal in which a selected target gene is stably silenced by utilizing the methods above.

Still other aspects of this invention involve producing a knock-out, non-human animal in which a selected target gene is stably silenced by utilizing the methods above, treating mammalian patients by silencing target genes that are associated with disease states, and manipulating stem cells to turn desired genes on and off at indicated intervals using the methods described herein.

These and other aspects of the invention will be apparent to one of skill in the art upon reading of the following detailed description of the invention.

#### DETAILED DESCRIPTION OF THE INVENTION

The present invention addresses the needs of the art by providing methods and compositions using a specifically defined inducible chimeric repressor fusion protein and nucleic acid molecules encoding that fusion protein to stably silence a target gene in cells. Among the embodiments of this invention, the inventors have created mammalian cell lines containing stably silenced target genes, which are useful in drug screening and development. The methods of development of such stable target gene suppression have other uses in the treatment of diseases, as well as the production of research tools and laboratory animals. As demonstrated in the examples below, the

methods and compositions of the invention can rapidly repress a target gene. Remarkably, in the absence of the ligand (inducer) that is initially employed to activate silencing of the target gene by the inventors' methods, the silent state of this transgene is maintained in clonal sub-populations after many cell divisions. This silenced state is epigenetically heritable, permitting many uses of cells manipulated to contain silenced exogenous or endogenous genes in biological research and in the treatment of diseases mediated by the expression of selected target genes.

*A. The Chimeric Repressor Fusion Protein*

A chimeric repressor fusion protein useful in the methods of this invention comprise minimally a first amino acid sequence comprising a Krüppel-Associated Box (KRAB) domain, fused to a second amino acid targeting sequence that binds to a target gene. These chimeric receptor proteins further contain the targeting sequence fused to a ligand-dependent binding domain of e.g., a nuclear hormone receptor that, in the presence of a ligand, permits the second amino acid sequence to bind to the target gene. The general structure of a suitable chimeric repressor is shown in Fig. 1A.

*1. KRAB Domain*

KRAB defines a highly conserved repression domain that is present in more than 220 human zinc-finger proteins (ZFP) that form the KRAB-ZFP superfamily of transcriptional silencers. A KRAB domain for use in this invention is the approximately 90 amino acid KRAB region from Kox1 gene, which binds the KAP1 protein and is a very strong, DNA-binding dependent repressor *in vivo* (see, e.g., J. R. Friedman *et al.*, 1996 *Genes Dev.*, 10:2067-2078; J. F. Margolin *et al.*, 1994 *Proc. Natl. Acad. Sci. USA*, 91:4509-4513; Moosemann *et al.*, 1977 *Biol. Chem.*, 378(7):669-677; International Patent Publication No. WO02/33104). Without wishing to be bound by theory, the inventors have surprisingly determined that this chimeric protein enables stable silencing of a transgene. This characteristic is apparently contributed by the presence of the KRAB domain. Other similar repressor domains have not demonstrated the ability to mediate stable silencing of a target gene when substituted for the KRAB domain in a similar chimeric fusion protein.

The KAP1 corepressor is a molecular scaffold protein that coordinates activities necessary for gene specific silencing. For example, KAP1 functions as a

corepressor by binding directly to specific promoters via the highly conserved KRAB repression domain present in more than 220 KRAB zinc-finger proteins in the human genome, suggesting that this mechanism is likely targeted to a large number of specific loci *in vivo* (M. Abrink *et al.*, 2001 *Proc. Natl. Acad. Sci. USA*, 98:1422-1426; H. Peng *et al.*, 2000 *J. Mol. Biol.*, 295:1139-1162). KAP1 also coordinates histone deacetylation via the recruitment of a unique form of the HDAC complex NuRD complex (Schultz *et al.*, 2001, *Genes Dev.*, 15:428-443) and histone H3 Lysine 9 methylation via the action of a novel KAP1-associated, SET domain protein with histone methyltransferase activity named SETDB1 (Schultz, D. C., *et al.*, 2002 *Genes Dev.*, 16, 919-932). SETDB1 demonstrates absolute specificity for histone H3-K9 methylation and enhances the recruitment of HP1 proteins to a euchromatic locus silenced by the chimeric fusion protein defined herein.

The HP1 family of heterochromatin proteins are small non-histone chromosomal proteins that are composed of a signature NH<sub>2</sub>-terminal chromodomain (CD), a COOH-terminal chromoshadow domain (CSD) and a variable hinge region that separates these two domains. The CD binds directly, with high affinity to the MeK9 residue in the histone H3 tail. The CSD is a homo-dimerization domain that directly recognizes a consensus pentapeptide sequence, PxVxL SEQ ID NO: 45 with high affinity, which is present in a growing number of nuclear proteins that may target the HP1 protein to specific genes or sub-nuclear compartments. Thus, the HP1 proteins appear to be bi-functional cross-linkers, which are anchored to chromatin, at least in part, via a highly specific modification of histone H3. KAP1 also directs the direct binding and deposition of HP1 mediated by the highly conserved PxVxL motif present in KAP1 (M. S. Lechner *et al.*, 2000 *Mol. Cell. Biol.*, 20:6449-6465) that interacts with the CSD of HP1.

Recent circumstantial evidence suggests a considerably broader role for HP1 and heterochromatin in the regulation of gene expression in mammals. First, the human and mouse genomes contain at least three HP1 homologs, and some of which display constitutively euchromatic localization. Second, there is the potential for hetero-dimerization among these HP1 isoforms, which could increase their target range for protein-protein interactions. Third, there are a host of regulatory proteins, which appear able to target HP1 to specific loci via a specific interaction with the

chromoshadow domain. Finally, there is distinct tissue-, cell-type specificity, and developmental regulation of the HP1 isoforms including an emerging role in cancer development and tumor progression. All these observations suggest that the mammalian HP1 proteins may participate in the regulation of specific transcriptomes *in vivo*.

The KAP1 corepressor is targeted by KRAB-ZFPs to specific loci and that can recruit and coordinate at such loci many of the components required for HP1-mediated gene silencing. KAP1 mediates direct deposition of HP1 protein and coordination of DNA methylation. These above-described actions of KAP1 collectively facilitate the nucleation of a localized heterochromatin environment to silence gene expression.

In addition to the use of known KRAB domain sequences in this chimeric protein, it is anticipated that suitable variants of the KRAB domain can be used in place of the naturally occurring sequence. A "variant" of a KRAB domain is any analogue, fragment, derivative, or mutant which is derived from a KRAB domain and which retains the ability to bind KAP1 and retain the DNA-dependent repressor activity characteristic of the wild-type KRAB domain. Different variants of the polypeptide or protein may exist in nature. These variants may be allelic variations characterized by differences in the nucleotide sequences of the structural gene coding for the protein, or may involve differential splicing or post-translational modification. The skilled artisan can produce variants having single or multiple amino acid substitutions, deletions, additions, or replacements. These variants may include, *inter alia*: (a) variants in which one or more amino acid residues are substituted with conservative or non-conservative amino acids, (b) variants in which one or more amino acids are added to the polypeptide or protein, (c) variants in which one or more of the amino acids includes a substituent group, and (d) variants in which the polypeptide or protein is fused with another polypeptide such as serum albumin. Techniques for obtaining these variants, including genetic (suppressions, deletions, mutations, etc.), chemical, and enzymatic techniques, are known to persons having ordinary skill in the art.

Modifications at the nucleic acid level of the KRAB domain may include, for example, modifications to the nucleotide sequences which are silent or

which change the amino acids, e.g. to improve expression. Also included are allelic variations, caused by the natural degeneracy of the genetic code. Additional homologous KRAB domain sequences can include mutants including 5' or 3' terminal deletion (truncation) or internal deletion, which deletion mutant sequences may be expressed for the purpose of affecting the activity of the full-length or wild-type KRAB or fragments.

Useful KRAB domains may be modified by conventional site-selected mutagenesis or other techniques. It is within the skill of the art to obtain or prepare synthetically or recombinantly a wide number of KRAB domain sequences, or modified polynucleotide sequences encoding same.

## 2. *Targeting Sequence and Target Gene*

Another portion of the chimeric protein is the targeting sequence that functions to bind the protein to the target gene to be silenced. In one embodiment, a single targeting sequence is present in each construct. In another embodiment, multiple targeting sequences may be present. In one embodiment, this targeting sequence is heterologous to the KRAB domain sequence. The targeting sequence is preferably a DNA binding domain (DBD) that binds to a DNA binding site in the target gene. This DNA binding site is preferably a monomeric, extended non-degenerate DNA sequence in the target gene. Preferably these binding sites are about 10-20 amino acids in length. A variety of DNA binding sites are known in the art for many target genes. A non-exclusive list of suitable binding sites includes paired box, homeodomain, basic-leucine zipper, basic-helix loop helix, winged helix, C<sub>4</sub>C<sub>3</sub> zinc finger, and C<sub>2</sub>H<sub>2</sub> zinc finger binding sites. Specific examples of DNA binding sites include PAX, HOX, MSX, POS, JUN, MYC, MYOD, FLI, FKHR, HNF3, NHR, KRAB, SNAG, WT-1, EGR, OLI, p53, and E2F sites. Still other sites useful as DNA binding sites are synthetic binding sites that are designed for specific binding to specific target genes. Such synthetic binding sites may be about 30 amino acids or greater in length. One of skill in the art may select any suitable binding site from available public and commercial sources for use in the chimeric repressor protein of this invention.

In another embodiment, this targeting sequence is a protein association motif (PAM). Exemplary PAMs include, without limitation, LEF1, KAP-t, CAF1,

RB, and NHR. One of skill in the art may select any suitable PAM from available public and commercial sources for use in the chimeric repressor protein of this invention.

5 The identity of the target gene to which the targeting sequence binds is determined by the reason for silencing, e.g., therapeutic, research, diagnostic, etc. Therefore, the target gene may include a gene that is exogenously introduced into a cell for expression and use as a reporter that generates a detectable signal. Such targets genes are useful in methods for drug screening or research. Examples of reporter genes include, without limitation, the genes encoding luciferase (Luc), green  
10 fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT),  $\beta$ -galactosidase (LacZ),  $\beta$ -glucuronidase (Gus),  $\beta$ -lactamase, alkaline phosphatase, thymidine kinase, and geneticin, hygromycin or purimycin resistance, as well as others well known in the art. These sequences, when associated with regulatory elements which drive their expression, provide signals detectable by conventional means. Such conventional  
15 means include, without limitation, enzymatic, radiographic, colorimetric, fluorescence or other spectrographic assays, fluorescent activated cell sorting assay and immunological assays, including ELISA, RIA and immunohistochemistry.

In another embodiment, the target gene is a gene endogenous to the cell, for example, a gene encoding a protein for which expression control, e.g.,  
20 suppression or sequential, controlled suppression and expression, is desirable. Such target genes are useful for research and therapy of disease, and can include, without limitation, genes that express growth factors, oncogenes, cytokines, chemokines and kinases, among others. Other useful gene products encoded by the target gene include hormones and growth and differentiation factors including, without limitation, insulin,  
25 glucagon, growth hormone (GH), parathyroid hormone (PTH), growth hormone releasing factor (GRF), follicle stimulating hormone (FSH), luteinizing hormone (LH), human chorionic gonadotropin (hCG), vascular endothelial growth factor (VEGF), angiopoietins, angiostatin, granulocyte colony stimulating factor (GCSF), erythropoietin (EPO), connective tissue growth factors (CTGF), basic fibroblast  
30 growth factor (bFGF), acidic fibroblast growth factor (aFGF), epidermal growth factor (EGF), transforming growth factor  $\alpha$  (TGF $\alpha$ ), platelet-derived growth factor (PDGF), insulin-like growth factors I and II (IGF-I and IGF-II), any one of the

transforming growth factor  $\beta$  (TGF $\beta$ ) superfamily comprising TGF $\beta$ , activins, inhibins, or any of the bone morphogenic proteins (BMP) BMPs 1-15, any one of the heregulin/neuregulin/ARIA/neu differentiation factor (NDF) family of growth factors, nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophins NT-3 and NT-4/5, ciliary neurotrophic factor (CNTF), glial cell line derived neurotrophic factor (GDNF), neurturin, agrin, any one of the family of semaphorins/collapsins, netrin-1 and netrin-2, hepatocyte growth factor (HGF), ephrins, noggin, sonic hedgehog and tyrosine hydroxylase.

Still other useful target genes encode proteins that regulate the immune system including, without limitation, cytokines and lymphokines such as thrombopoietin (TPO), interleukins (IL) IL-1 $\alpha$ , IL-1 $\beta$ , IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, and IL-17, monocyte chemoattractant protein (MCP-1), leukemia inhibitory factor (LIF), granulocyte-macrophage colony stimulating factor (GM-CSF), Fas ligand, tumor necrosis factors  $\alpha$  and  $\beta$  (TNF $\alpha$  and TNF $\beta$ ), interferons (IFN) IFN- $\alpha$ , IFN- $\beta$  and IFN- $\gamma$ , stem cell factor, and flk-2/flt3 ligand. Target genes also may encode products produced by the immune system, including, without limitations, immunoglobulins, T cell receptors, class I and class II MHC molecules, complement regulatory proteins such as membrane cofactor protein (MCP), decay accelerating factor (DAF), CR1, CR2 and CD59.

Yet other useful target genes encode a receptor for a protein selected from among hormones, growth factors, cytokines, lymphokines, regulatory proteins and immune system proteins. Target genes include the encoded receptors for cholesterol regulation, including the LDL receptor, HDL receptor, VLDL receptor, and the scavenger receptor, glucocorticoid receptors, estrogen receptors, Vitamin D receptors. Other useful target genes encode products including transcription factors such as *jun*, *fos*, *max*, *mad*, serum response factor (SRF), AP-1, AP-2, *myb*, MRG1, CREM, Alx4, FREAC1, NF- $\kappa$ B, members of the leucine zipper family, C2H4 zinc finger proteins, including Zif268, EGR1, EGR2, C6 zinc finger proteins, including the glucocorticoid and estrogen receptors, POU domain proteins, exemplified by Pit1, homeodomain proteins, including HOX-1, basic helix-loop-helix proteins, including *myc*, MyoD and myogenin, ETS-box containing proteins, TFE3, E2F, ATF1, ATF2,

ATF3, ATF4, ZF5, NFAT, CREB, HNF-4, C/EBP, SP1, CCAAT-box binding proteins, interferon regulation factor 1 (IRF-1), Wilms tumor protein, ETS-binding protein, STAT, GATA-box binding proteins, e.g., GATA-3, and the forkhead family of winged helix proteins.

5                    Still other useful target genes encode carbamoyl synthetase I, ornithine transcarbamylase, arginosuccinate synthetase, arginosuccinate lyase, arginase, fumarylacetoacetate hydrolase, phenylalanine hydroxylase, alpha-1 antitrypsin, glucose-6-phosphatase, low-density-lipoprotein receptor, porphobilinogen deaminase, factor VIII, factor IX, cystathione beta-synthase, branched chain ketoacid  
10    decarboxylase, albumin, isovaleryl-CoA dehydrogenase, propionyl CoA carboxylase, methyl malonyl CoA mutase, glutaryl CoA dehydrogenase, insulin, beta-glucosidase, pyruvate carboxylase, hepatic phosphorylase, phosphorylase kinase, glycine decarboxylase (also referred to as P-protein), H-protein, T-protein, Menkes disease protein, tumor suppressors (e.g., p53), cystic fibrosis transmembrane regulator  
15    (CFTR), and the product of Wilson's disease gene PWD.

                    The selection of the targeting sequence and target gene from among many known and available sequences, as well as the many targeting sequences and genes yet to be identified, is within the ability of the person of skill in the art. The particular identity of these components of the fusion protein is not a limitation of this  
20    invention.

### 3.        *Switching Component, e.g., Ligand Binding Domain*

                    Another component of the fusion protein is a component that functions as a "switch" to turn on the binding ability of the targeting sequence in the presence of a ligand or inducer, and make the binding inactive in the absence of the inducer.  
25    Many components suitable for this purpose are known in the art, such as inducible promoters, including without limitation, those listed below, and other "gene switch" components, such as those described in e.g., US Patent Nos. 6,479,653; 6,379,945; 6,258,603; 6,339,070; 6,380,373; 6,504,082; among many others.

                    In one exemplified embodiment, a ligand binding domain (LBD) from  
30    a ligand-dependent (ligand-inducible) nuclear receptor is used as this switch. In one embodiment, the LBD is fused to the targeting sequence. Numerous suitable nuclear receptors and their LBDs and ligands are known in the art (see, e.g., R. M. Evans,



1988 *Science*, 240:889). The selections of particular LBD, its nuclear receptor, and ligand are not limitations of this invention. For example, steroid hormone receptors are members of the nuclear receptor superfamily and are found in vertebrate and invertebrate cells. See, e.g., International Patent Publication No. WO01/09180, published February 8, 2001, which discusses biological receptors modified to have novel specificities for xenobiotics and other non-native ligands. Members of the nuclear receptor superfamily include, without limitation, a modified or native steroid/thyroid nuclear receptor superfamily protein, such as the ecdysone (see Yao, T.P. *et al* 1993 *Nature*, 366: 476-479; Yao, T.-P. *et al*, 1992 *Cell*, 71: 63-72), the estrogen, retinoid X, progesterone, glucocorticoid, vitamin D, retinoic acid, and peroxisome proliferation receptor proteins.

Other known nuclear receptors include, without limitation, thyroid hormone receptor  $\alpha$  (TR $\alpha$ ), thyroid receptor 1 (c-erbA-1), thyroid hormone receptor  $\alpha$  (THRA), thyroid hormone receptor  $\beta$  (TR $\beta$ ), thyroid hormone receptor  $\beta$  (THRB), retinoic acid receptor  $\alpha$  (RAR $\alpha$ ), retinoic acid receptor  $\beta$  (RAR $\beta$ ), hepatoma (HAP), retinoic acid receptor  $\gamma$  (RAR $\gamma$ ), retinoic acid receptor gamma-like (RARD), peroxisome proliferator-activated receptor  $\alpha$  (PPAR $\alpha$ ), peroxisome proliferator-activated receptor  $\beta$  (PPAR $\beta$ ), peroxisome proliferator-activator related receptor (NUC-1), peroxisome proliferator-activated receptor  $\delta$  (PPAR $\delta$ ), peroxisome proliferator-activator related receptor (FFAR), peroxisome proliferator-activated receptor  $\gamma$  (PPAR $\gamma$ ), orphan receptor encoded by non-encoding strand of thyroid hormone receptor  $\alpha$  (REVERB $\alpha$ ), v-erb A related receptor (EAR-1), v-erb related receptor (EAR-1A), orphan receptor encoded by non-encoding strand of thyroid hormone receptor  $\beta$  (REVERB $\beta$ ), v-erb related receptor (EAR-1 $\beta$ ), orphan nuclear receptor BD73 (BD73), rev-erb A-related receptor (RVR), zinc finger protein 126 (HZF2), ecdysone-inducible protein E75 (E75), ecdysone-inducible protein E78 (E78), *Drosophila* receptor 78 (DR-78), retinoid-related orphan receptor  $\alpha$  (ROR $\alpha$ ), retinoid Z receptor  $\alpha$  (RZR $\alpha$ ), retinoid related orphan receptor  $\beta$  (ROR $\beta$ ), retinoid Z receptor  $\beta$  (RZR $\beta$ ), retinoid-related orphan receptor  $\gamma$  (ROR $\gamma$ ), retinoid Z receptor  $\gamma$  (RZR $\gamma$ ), retinoid-related orphan receptor (TOR), hormone receptor 3 (HR-3), *Drosophila* hormone receptor 3 (DHR-3), myohemerythin (MHR-3), growth hormone

receptor 3 (GHR-3), *C. elegans* nuclear receptor 3 (CNR-3), *C. elegans* hormone  
 receptor 3 (CHR-3), *C. elegans* nuclear receptor 14 (CNR-14), ecdysone receptor  
 (ECR), ubiquitous receptor (UR), orphan nuclear receptor (OR-1), NER-1, receptor-  
 interacting protein 15 (RIP-15), liver X receptor  $\beta$  (LXR $\beta$ ), steroid hormone receptor  
 5 like protein (RLD-1), liver X receptor (LXR), liver X receptor  $\alpha$  (LXR $\alpha$ ), farnesoid  
 X receptor (FXR), receptor-interacting protein 14 (RIP-14), HRR-1, vitamin D  
 receptor (VDR), orphan nuclear receptor (ONR-1), pregnane X receptor (PXR),  
 steroid and xenobiotic receptor (SXR), benzoate X receptor (BXR), nuclear receptor  
 (MB-67), constitutive androstane receptor 1 (CAR-1), constitutive androstane  
 10 receptor  $\alpha$  (CAR $\alpha$ ), constitutive androstane receptor 2 (CAR-2), constitutive  
 androstane receptor  $\beta$  (CAR $\beta$ ), *Drosophila* hormone receptor 96 (DHR-96), nuclear  
 hormone receptor 1 (NHR-1), hepatocyte nuclear factor 4 (HNF-4), hepatocyte  
 nuclear factor 4G (HNF-4G), hepatocyte nuclear factor 4B (HNF-4B), DHNF-4,  
 hepatocyte nuclear factor 4D (HNF-4D), retinoid X receptor  $\alpha$  (RXR $\alpha$ ), retinoid X  
 15 receptor  $\beta$  (RXR $\beta$ ), H-2 region II binding protein (H-2RIIBP), nuclear receptor co-  
 regulator-1 (RCoR-1), retinoid X receptor  $\gamma$  (RXR $\gamma$ ), Ultraspiracle (USP), 2C1,  
 chorion factor 1 (CF-1), testicular receptor (TR-2), testicular receptor (TR2-11), TR4,  
 TAK-1, *Drosophila* hormone receptor (DHR78), Tailless (TLL), tailless homolog  
 (TLX), XTLL, chicken ovalbumin upstream promoter transcription factor I (COUP-  
 20 TFI), chicken ovalbumin upstream promoter transcription factor A (COUP-TFA),  
 EAR-3, SVP-44, chicken ovalbumin upstream promoter transcription factor II  
 (COUP-TFII), chicken ovalbumin upstream promoter transcription factor B (COUP-  
 TFB), ARP-1, SVP-40, SVP, chicken ovalbumin upstream promoter transcription  
 factor III (COUP-TFIII), chicken ovalbumin upstream promoter transcription factor G  
 25 (COUP-TFG), SVP-46, EAR-2, estrogen receptor  $\alpha$  (ER $\alpha$ ), estrogen receptor  $\beta$   
 (ER $\beta$ ), estrogen related receptor 1 (ERR1), estrogen related receptor  $\alpha$  (ERR $\alpha$ ),  
 estrogen related receptor 2 (ERR2), estrogen related receptor  $\beta$  (ERR $\beta$ ),  
 glucocorticoid receptor (GR), mineralocorticoid receptor (MR), progesterone receptor  
 (PR), androgen receptor (AR), nerve growth factor induced gene B (NGFI-B), nuclear  
 30 receptor similar to Nur-77 (TRS), N10, Orphan receptor (NUR-77), human early  
 response gene (NAK-1), Nurr related factor 1 (NURR-1), a human immediate-early

response gene (NOT), regenerating liver nuclear receptor 1 (RNR-1), hematopoietic zinc finger 3 (HZF-3), Nur related protein -1 (TINOR), nuclear orphan receptor 1 (NOR-1), NOR1 related receptor (MINOR), *Drosophila* hormone receptor 38 (DHR-38), *C. elegans* nuclear receptor 8 (CNR-8), C48D5, steroidogenic factor 1 (SF1), endozepine-like peptide (ELP), fushi tarazu factor 1 (FTZ-F1), adrenal 4 binding protein (AD4BP), liver receptor homolog (LRH-1), Ftz-F1-related orphan receptor A (xFFrA), Ftz-F1-related orphan receptor B (xFFrB), nuclear receptor related to LRH-1 (FFLR), nuclear receptor related to LRH-1 (PHR), fetoprotein transcription factor (FTF), germ cell nuclear factor (GCNFM), retinoid receptor-related testis-associated receptor (RTR), knirps (KNI), knirps related (KNRL), Embryonic gonad (EGON) receptor, *Drosophila* gene for ligand dependent nuclear receptor (EAGLE), nuclear receptor similar to trithorax (ODR7), trithorax, dosage sensitive sex reversal adrenal hypoplasia congenital critical region chromosome X gene (DAX-1), adrenal hypoplasia congenita and hypogonadotropic hypogonadism (AHCH), and short heterodimer partner (SHP).

Other components of suitable inducible systems are available from, e.g., Invitrogen, Clontech and Ariad Corporations. One of skill in the art may readily select suitable sequences from among those known. However, in one embodiment, the nuclear receptor is a hormone-inducible nuclear receptor, such as the estrogen receptor or the progesterone receptor.

The identity of the ligand, therefore, depends upon the particular inducible "switch" or LBD used in the chimeric fusion protein useful in this invention. For example, in the case of the estrogen receptor, a suitable ligand is the hormone 4- hydroxytamoxifen (4-OHT). In the case of the progesterone receptor, a suitable ligand is medroxyprogesterone acetate or levonorgestral or progesterone. Ligands for the above-noted receptors are well-known and within the knowledge of the person of skill in the art.

#### *B. The Nucleic Acid Molecule Encoding the Chimeric Fusion Protein*

Because the chimeric fusion protein must be delivered or introduced to a cell in the methods discussed herein, a nucleic molecule encoding the fusion protein is employed in the methods of this invention. In addition to DNA, and occasionally RNA sequences encoding the components of the fusion protein described above, the

nucleic acid molecule also contains regulatory sequences capable of directing expression of the fusion protein in a parent cell.

5 The nucleic acid molecule which encodes the fusion protein and introduces it into the parent cell is a recombinant vector, and includes both viral vectors and non-viral vectors (including non-viral methods of delivery of a nucleic acid molecule into a cell).

10 A variety of viral vector systems are known in the art. Examples of such vectors include recombinant adenoviral vectors, herpes simplex virus (HSV)-based vectors, adeno-associated viral (AAV) vectors, hybrid adenoviral/AAV vectors, recombinant retroviruses or lentiviruses which are constructed to carry or express a selected nucleic acid composition of interest.

15 Retrovirus vectors that can be employed include those described in EP 0 415 731; International Patent Publication Nos. WO 90/07936; WO 94/03622; WO 93/25698; and WO 93/25234; U.S. Pat. No. 5, 219,740; International Patent Publication Nos. WO 93/11230 and WO 93/10218; Vile and Hart, 1993 *Cancer Res.* 53:3860-3864; Vile and Hart, 1993 *Cancer Res.* 53:962-967; Ram *et al.*, 1993 *Cancer Res.* 53:83-88; Takamiya *et al.*, 1992 *J. Neurosci. Res.* 33:493-503; Baba *et al.*, 1993 *J. Neurosurg.* 79:729-735; US Patent No. 4,777,127; GB Patent No. 2,200,651; and EP 0 345 242. Examples of suitable recombinant retroviruses include those described  
20 in International Patent Publication No. WO 91/02805.

25 Alphavirus-based vectors may also be used as the nucleic acid molecule encoding the chimeric fusion protein. Such vectors can be constructed from a wide variety of alphaviruses, including, for example, Sindbis virus vectors, Semliki forest virus (ATCC VR-67; ATCC VR-1247), Ross River virus (ATCC VR-373; ATCC VR-1246) and Venezuelan equine encephalitis virus (ATCC VR-923; ATCC VR-1250; ATCC VR 1249; ATCC VR-532). Representative examples of such vector systems include those described in U.S. Pat. Nos. 5,091,309; 5,217,879; and 5,185,440; and International Patent Publication Nos. WO 92/10578; WO 94/21792; WO 95/27069; WO 95/27044; and WO 95/07994.

30 Examples of adenoviral vectors include those described by Berkner, 1988 *Biotechniques* 6:616-627; Rosenfeld *et al.*, 1991 *Science* 252:431-434; International Patent Publication No. WO 93/19191; Kolls *et al.*, 1994 *PNAS*

91:215-219; Kass-Eisler *et al.*, 1993 *PNAS* 90:11498-11502; Guzman *et al.*, 1993 *Circulation* 88:2838-2848; Guzman *et al.*, 1993 *Cir. Res.* 73:1202-1207; Zabner *et al.*, 1993 *Cell* 75:207-216; Li *et al.*, 1993 *Hum. Gene Ther.* 4:403-409; Cailaud *et al.*, 1993 *Eur. J. Neurosci.* 5:1287-1291; Vincent *et al.*, 1993 *Nat. Genet.* 5:130-134; Jaffe *et al.*, 1992 *Nat. Genet.* 1:372-378; and Levrero *et al.*, 1991 *Gene* 101:195-202.

Exemplary adenovirus vectors include those described in International Patent Publication Nos. WO 94/12649; WO 93/03769; WO 93/19191; WO 94/28938; WO 95/11984 and WO 95/00655. Other adenovirus vectors include those derived from chimpanzee adenoviruses, such as those described in US Patent No. 6,083,716.

Another viral vector is based on a parvovirus such as an adeno-associated virus (AAV). Representative examples include the AAV vectors disclosed by Srivastava in International Patent Publication No. WO 93/09239, Samulski *et al.*, 1989 *J. Virol.* 63:3822-3828; Mendelson *et al.*, 1988 *Virol.* 166:154-165; and Flotte *et al.*, 1993 *PNAS* 90:10613-10617. Other particularly desirable AAV vectors include those based upon AAV1; see, International Patent Publication No. WO 00/28061, published May 18, 2000. Other desirable AAV vectors include those which are pseudotyped, i.e., contain a minigene composed of AAV 5' ITRS, a transgene, and AAV 3' ITRs packaged in a capsid of an AAV serotype heterologous to the AAV ITRs. Methods of producing such pseudotyped AAV vectors are described in detail in International Patent Publication No. WO01/83692.

The nucleic acid molecule of the invention may also include non-viral vectors or methods for delivery of the sequence encoding the chimeric repressor protein to the parent cell according to this invention. A variety of non-viral vectors are known in the art and may include, without limitation, plasmids, "naked" DNA and DNA condensed with cationic lipids or polymers. In one embodiment, the polymers may include traditional polymers and non-traditional polymers such as cyclodextrin-containing polymers and protective, interactive noncondensing polymers, among others. The "naked" DNA and DNA condensed with cationic lipids or polymers are typically delivered to the cells using chemical methods. A number of chemical methods are known in the art for cell delivery and include using lipids, polymers, or proteins to complex with DNA, optionally condensing the same into particles, and

delivering to the cells. Another non-viral chemical method includes using cations to condense DNA, which is then placed in a liposome and used according to the present invention. See, C. Henry, 2001 *Chemical and Engineering News*, 79(48):35-41.

Whether the nucleic acid molecule is a viral vector or non-viral vector,  
5 it may optionally contain regulatory sequences in addition to the sequences encoding the fusion protein. For example, such regulatory sequences comprise a promoter which drives expression of the fusion protein. Suitable promoters may be readily selected from among constitutive promoters, inducible promoters, tissue-specific promoters and others.

10 Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV), LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) (see, e.g., Boshart *et al*, 1985 *Cell*, 41:521-530), the SV40 promoter, the dihydrofolate reductase promoter, the  $\beta$ -actin promoter, the phosphoglycerol kinase (PGK)  
15 promoter, and the EF1 $\alpha$  promoter (Invitrogen).

Inducible promoters are regulated by exogenously supplied compounds, including, the zinc-inducible sheep metallothioneine (MT) promoter, the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter, the T7 polymerase promoter system (WO 98/10088); the ecdysone insect promoter (No  
20 *et al*, 1996 *Proc. Natl. Acad. Sci. USA*, 93:3346-3351), the tetracycline-repressible system (Gossen *et al*, 1992 *Proc. Natl. Acad. Sci. USA*, 89:5547-5551), the tetracycline-inducible system (Gossen *et al*, 1995 *Science*, 268:1766-1769, see also Harvey *et al*, 1998 *Curr. Opin. Chem. Biol.*, 2:512-518), the RU486-inducible system (Wang *et al*, 1997 *Nat. Biotech.*, 15:239-243 and Wang *et al*, 1997 *Gene Ther.*, 4:432-  
25 441) and the rapamycin-inducible system (Magari *et al*, 1997 *J. Clin. Invest.*, 100: 2865-2872).

Useful tissue-specific promoters include the promoters from genes encoding skeletal  $\beta$ -actin, myosin light chain 2A, dystrophin, muscle creatine kinase, as well as synthetic muscle promoters with activities higher than naturally-occurring  
30 promoters (see Li *et al.*, 1999 *Nat. Biotech.*, 17:241-245). Examples of promoters that are tissue-specific are known for liver (albumin, Miyatake *et al.* 1997 *J. Virol.*, 71:5124-32; hepatitis B virus core promoter, Sandig *et al.*, 1996 *Gene Ther.*, 3:

1002-9; alpha-fetoprotein (AFP), Arbuthnot *et al.*, 1996 *Hum. Gene Ther.*,  
7:1503-14), bone (osteocalcin, Stein *et al.*, 1997 *Mol. Biol. Rep.*, 24:185-96; bone  
sialoprotein, Chen *et al.*, 1996 *J. Bone Miner. Res.*, 11:654-64), lymphocytes (CD2,  
Hansal *et al.*, 1988 *J. Immunol.*, 161:1063-8; immunoglobulin heavy chain; T cell  
5 receptor  $\alpha$  chain), neuronal (neuron-specific enolase (NSE) promoter, Andersen *et al.*  
1993 *Cell. Mol. Neurobiol.*, 13:503-15; neurofilament light-chain gene, Piccioli *et al.*,  
1991 *Proc. Natl. Acad. Sci. USA*, 88:5611-5; the neuron-specific vgf gene, Piccioli *et al.*,  
1995 *Neuron*, 15:373-84); among others.

See, e.g., International Patent Publication No. WO00/55335 for  
10 additional lists of known promoters useful in this context.

Other regulatory sequences that may be present in a nucleic acid  
molecule of this invention include, among others, epitope tags, nuclear localization  
sequences, IRES elements, TATA elements, polyadenylation sites, restriction enzyme  
cleavage sites, selectable markers and the like. Selection of promoters and other  
15 common vector elements are conventional and many such sequences are available  
(see, e.g., Sambrook *et al.*, *Molecular Cloning. A Laboratory Manual*, Cold Spring  
Harbor Laboratory, New York, 1989 and references cited therein at, for example,  
pages 3.18-3.26 and 16.17-16.27 and Ausubel *et al.*, *Current Protocols in Molecular  
Biology*, John Wiley & Sons, New York, 1989).

20 These nucleic acid molecules are introduced into parent cells. By the  
term "parent cell" is meant the cell in which the nucleic molecule is deliberately  
introduced. Thereafter, cells resulting from normal population doublings of the parent  
cell containing the nucleic acid molecule are referred to herein as "progeny cells".

Depending upon the use to which the compositions and methods of this invention are  
25 placed, the cell may be selected from any biological organism, including prokaryotic  
(e.g., bacterial) cells and eukaryotic cells, including, insect cells, yeast cells. The cells  
employed in the various methods and compositions of this invention are mammalian  
cells. Cells are selected from among any mammalian species, such as human,  
monkey, mouse, rat, rabbit, and hamster, among others. In one embodiment, the cells  
30 employed in the various methods and compositions of this invention are human cells.

Classes of mammalian cells include multipotent stem cells, pluripotent stem  
cells, and cells of any tissue that contain a target gene for which expression is desired

to be silenced. Exemplary cells for use in this invention include specifically, without limitation, primary fibroblast, hepatocyte and myoblast cells derived from mammals cells; various murine cells, e.g., 10T1/2 and WEHI cells, African green monkey cells such as VERO, COS1, COS7, BSC1, BSC 40, and BMT 10, and human cells such as  
5 WI38, MRC5, A549, human embryonic retinoblast (HER), human embryonic kidney (HEK), human embryonic lung (HEL), and TH1080 cells. Other appropriate cells include 293 cells (human embryonic kidney cells which express adenoviral E1a and E1b proteins), 911, PER.C6 cells (human embryonic retinoblast cells that express adenoviral E1; see WO 97/19463), GH329 cells (a cell line that expresses adenoviral  
10 E1); 27-18 cells, 84-31 cells (293-based cells that express adenovirus E1a, E1b and E4 (G. Gao, 1996 J. Virol., 70(12):8934-8943), 10-3 cells (293-based cells that express adenovirus E1a, E1b and E4ORF6 (G. Gao, 1996 J. Virol), 3T3 cells (mouse embryonic fibroblast cell line), NIH3T3 cells (subline of 3T3 cells), HepG2 cells (human liver carcinoma cell line), Saos-2 cells (human osteogenic sarcomas cell line),  
15 HuH7 cells or HeLa cells (human carcinoma cell line). Still other suitable mammalian cells include 10T1/2, BHK, MDCK, Saos, C2C12, L cells, HT1080, CHO, and BKH.

The chimeric repressor fusion protein, its various components parts and the nucleic acid molecules described above may be constructed recombinantly  
20 using conventional molecular biology techniques, site-directed mutagenesis, genetic engineering or PCR, and the like by utilizing the information provided herein. For example, methods for producing the above-identified modifications of the sequences include mutagenesis of certain nucleotides and/or insertion or deletion of nucleotides, or codons, thereby effecting the polypeptide sequence by insertion or deletion of, e.g.,  
25 non-natural amino acids. Such methods are known and may be selected by one of skill in the art. Similarly, methods for producing plasmid, other non-viral vector constructs or viral vector constructs encoding the chimeric protein, and/or any other reporter molecules used herein are well-known in the art, as are methods for using expression systems to produce the fusion protein.

30



*C. A Cell Containing a Stably Silenced Target Gene and Methods of Production*

One embodiment of the present invention is a method for producing a cell, preferably a cloned cell, containing an epigenetically heritable, stably silenced, target gene and the cell itself. This method involves introducing into a parent cell a nucleic acid molecule expressing an above-defined chimeric repressor fusion protein. Using the teachings described herein, the elements of the nucleic acid molecule encoding the chimeric fusion protein and the identity of the parent cell may all be selected by one of skill in the art. By "introducing" the nucleic acid into the cell is meant delivering the nucleic acid molecules to the cells in any manner known to one in the art, including, without limitation, transfection, infection, electroporation, sonoporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion, or particle bombardment. However, other methods known by those skilled in the art may be utilized. In one embodiment, the nucleic acid molecules may be transfected into the host cell and exist stably in the cell as an episome. In another embodiment, the sequences encoding the chimeric fusion protein are stably integrated into the genome of the cell. Another embodiment has the sequences transiently expressed in the host cell.

The parent cell is then cultured under known conditions suitable for that cell type. In one embodiment, the cell is cultured in the presence of an effective amount of the appropriate ligand that induces the LBD of the chimeric protein. The "effective amount" of ligand utilized according to the present invention is typically that which is necessary to permit the targeting amino acid sequence of the chimeric protein to bind to the target gene. An effective amount of the ligand ranges from about 1 nM to about 1000 nM of ligand. In one embodiment, the effective amount of ligand is about 25 nM to about 750 mM. In another embodiment, the effective amount of ligand is about 100 to about 600 mM. In still other embodiments, the effective amount of ligand is about 500 mM. For example, wherein the LBD is from the estrogen receptor, the appropriate ligand is 4-OHT, as used in the examples below. In the examples below, an effective amount of the ligand 4-OHT is about 500 mM.

The parent cell culture is cultured in the ligand for a sufficient time to repress expression of the target gene. Without wishing to be bound by theory, the

inventors believe that during this time, the KRAB domain mediates the development of a complex comprising the KAP1 protein, the HP1 protein and the SETDB1B enzymes at the site of the target gene, which in turn suppresses the expression of the target gene. In one embodiment, therefore, the “sufficient time” for this repression is defined as culturing the cells in the presence of the ligand for greater than 12 hours. In still another embodiment, the time for the cell to be exposed to ligand prior to triggering repressor activity is between about 24 hours to about 150 hours. In still another embodiment the time for exposure to ligand is greater than 48 hours. In another embodiment the time for exposure is greater than 90 hours. It has been observed that the longer the parent cells are exposed to the ligand, the greater is the resulting proportion of progeny cells with stably silenced target genes. For example, the examples below employ a ligand exposure time of about 96 hours.

Once the cell culture has been exposed to the ligand, these treated cells are preferably subjected to conventional single cell cloning in the absence of ligand to obtain single subclones thereof. Alternatively, progeny cells with stably silenced target genes in the absence of ligand/inducer are not cloned. In one embodiment, single cell clones are cultured for at least about 25 cell doublings and maintain a silenced target gene. Preferably the clones or cells are cultured for at least about 40 cell doublings. In yet another embodiment, the cells are cultured for about 50 cell doublings and maintain the silenced target gene. Thereafter the clones and/or cells are assayed for expression of the target gene, using any conventional assay that is typical for the selected target gene. Non-expressing cells are isolated from expressing cells. The isolated, non-expressing cells or clones contain a target gene that is stably silenced and that can transfer this silenced state of the gene epigenetically to its progeny. Examples of such clones or cells of the invention containing a stably silenced target gene are described below.

Such subclones or other cells may contain stably silenced endogenous target genes, which may be useful in the drug screening and research into diseases. Alternatively or additionally, such subclones or cells may contain stably silenced reporter genes. The cloned cells in the examples below employ a reporter plasmid in the cell to enable the use of a reporter gene expressing luciferase as the target gene. See, for example, the schematic of Fig. 2A. In such embodiments, the method

described above further comprises a step of introducing into the cell a reporter vector containing a reporter target gene under the control of suitable regulatory sequences, preceded by multiple copies of a DNA-binding sequence capable of binding to the targeting sequence of the chimeric fusion protein, and a selectable marker.

5                    Still another embodiment of the cells and subcloned cells of this invention is prepared by introducing to the cell multiple different nucleic acid molecules encoding multiple different chimeric fusion proteins in order to silence multiple different target genes in a single cell. The parent cell, according to the methods described above, thus is treated to contain a fusion protein to silence a first  
10                   gene, e.g., a reporter gene, and a different fusion protein to silence a second gene, e.g., an endogenous gene. In such a cell, the fusion proteins must differ in targeting sequences as well as contain different LBDs so as to be inducible with different ligands. Thus the above method would comprise introducing at least two different nucleic acid molecules into the parent cells and culturing the cells in effective  
15                   amounts of at least two different ligands. The ligand treatments may occur sequentially or simultaneously. In one embodiment, the cells may be exposed to the first ligand and then optionally singly cloned, followed by exposure of the clones to the second ligand for a suitable ligand exposure time. The resulting cells or subcloned single cells are allowed to double as described above. In yet another  
20                   embodiment the cells are cultured in each ligand individually and sequentially, prior to culturing or single cell cloning. The cells or clones are tested for silenced expression of one or both target genes by appropriate assays for the products of target gene expression. The resulting progeny cell or clone are stably silenced in at least two target genes. One of skill in the art may expand the methods described herein to  
25                   silence additional target genes in a single cell, if desired, and obtain the resulting cloned cell by following the teachings herein.

                    KRAB-KAP1 system is used herein as a component of a hormone inducible system in a mammalian cell line that allows transient and reversible targeting of endogenous KAP1 and HP1 and its associated activities to a highly  
30                   transcribed euchromatic reporter transgene. This target gene is rapidly repressed and adopts a highly localized compact chromatin structure that is enriched in HP1 and H3-MeK9. Remarkably, in the absence of ligand/hormone the silent state of this

transgene is maintained in clonal sub-populations after many cell divisions, an effect highly reminiscent of variegated gene expression observed in flies. The promoter region in silent clones is enriched in SETDB1, HP1 and H3-MeK9. Moreover, the CpG residues in the promoter region are hypermethylated in silent clones, compared to expressed clones, thus suggesting a link between establishing the H3-MeK9 mark and DNA methylation. KAP1 coordinates the establishment of highly localized heterochromatin-like silenced states at euchromatic genes and that these states are epigenetically heritable. Most importantly, KRAB-mediated repression of an endogenous target gene mirrors the physical characteristics observed for the engineered KRAB repressor system.

*D. Methods for Screening Drug Candidates*

In another embodiment, the invention provides a method for identifying a test molecule that activates or re-activates the expression of a stably silenced target gene. This method employs a progeny cell or cloned cell as described above, in which at least one target gene has been silenced. Desirably, the target gene (or at least one target gene if more than one target gene is silenced in the cell) is a reporter gene. In one embodiment, the cloned cell contains a nucleic acid encoding the chimeric repressor fusion protein as well as a nucleic acid molecule encoding a reporter target gene.

According to this method, a test compound or molecule is brought into contact with such a cloned cell or progeny thereof, such as by addition of varying amounts of the compound to the cell culture media. The cells or progeny thereof are left in contact with the test molecule, preferably for varying time periods. Such time periods can include periods of between about 1 hour to about 150 hours. One embodiment of suitable time periods for contact between the test molecule and the culture is between about 25 to about 100 hours. Another embodiment of suitable time periods for contact between the test molecule and the culture is between about 50 to about 75 hours. One of skill in the art may readily select the suitable contact time period based upon individual responses of the test molecules and controls.

Thereafter each sample of contacted cultured cells or clones is assayed by a conventional assay suitable for the product of the target gene. These assays detect expression of the product encoded by the target gene in the cultures contacted

with test molecule compared to cultures contacted with a control. For example, if the target gene is a fluorescent protein, the assay is simply the detection of fluorescence. If the target gene produces an enzyme, a suitable assay involves the addition of the substrate and observation of enzymatic activity. As stated above, selection of assays  
5 suitable to monitor expression of the product of the target gene may be readily selected by one of skill in the art, and is not a limitation of this invention. Detection of expression of the target gene following treatment of the cloned cell with the test molecule indicates that the test molecule reactivates expression of the silenced gene.

In one embodiment, the test molecule is one that effects or inhibits  
10 histone deacetylases, or histone methyltransferases, or DNA methyltransferases. The test molecule can effect or inhibit interactions between KRAB and KAP1, between KAP1 and HP1, between KAP1 and SETDB1 or between KAP1 and Mi-2 $\alpha$  at the target gene, i.e., it disrupts the binding of the components of the KAP1-HP1-SETDB1 complex at the target gene. Still other targets may be identified that impact other  
15 interactions related to silencing.

#### *E. Methods for Manipulating Gene Expression*

In yet a further aspect of this invention, a method is provided for manipulating expression of a target gene in a cell. Such manipulation may occur *in vitro*, e.g., in cell cultures. Alternatively, such manipulation may occur *ex vivo*, in  
20 cells removed from a patient's body and returned thereto after treatment. In still another embodiment of this method, the manipulation may occur by applying this method to cells *in vitro*, such as for gene therapy treatment.

The method comprises introducing to a cell containing a selected target gene an effective amount of a composition comprising a nucleic acid molecule  
25 expressing a chimeric repressor fusion protein, as defined above. The target gene, e.g. an endogenous or exogenous gene, in the cells now containing the nucleic acid molecule is stably silenced by transiently contacting the cell with an effective amount of a ligand. The ligand induces the LBD of the fusion protein for a sufficient time to stably suppress expression of the target gene in subsequent progeny of the cell  
30 containing the nucleic acid molecule in the absence of the ligand. For *in vitro* and *ex vivo* performance of this method, the effective amounts and sufficient times are as described above.

The method may further comprise the step of reactivating the silenced gene in the cells in which the above steps have suppressed expression of the target gene. According to this embodiment of the method, the cells containing the silenced target gene are contacted with a molecule or compound, such as those identified by the drug screening method described above. These compounds include, without limitation, compounds that disrupt the binding of the complex at the target gene, effect DNA methylation of the target gene or effect histone acetylation. It is anticipated that the effective amounts and times of administration of such a reactivating molecule or compound will be readily determined by one of skill in the art.

The ability of this method to stably silence and reactivate expression of a target gene at specified times by use of the ligands and test molecules described herein is valuable in research, in the establishment of knock-out laboratory animals, in drug screening and in the development of stem cells. Such methods are also useful in the treatment of diseases caused by overexpression or aberrant expression of target genes in a mammalian patient. The ability to stably silence a gene which assists in the initiation or development of cancers, immune disorders, etc., provides valuable treatment methods for such diseases. Similarly, the ability to suppress expression of a target gene and reactivate its expression later in time also has value in therapy of gene-related disorders of mammalian patients.

For performance of this method *in vivo* for gene therapy treatment of a disease or for the establishment of knock-out laboratory animals, where the cell is present in a living mammal, the methods involve administering an effective amount of a pharmaceutical composition comprising a nucleic acid molecule as described above. The introducing step involves administering the composition to the cell in a mammalian patient by a suitable route of administration. In one embodiment, a suitable route of administration includes direct or local delivery to the organ, tissue or cells containing the target gene. Other suitable routes of administration may be used, including, without limitation, parenteral administration, intraperitoneal administration, intravenous administration, intramuscular administration, subcutaneous administration, intradermal administration, intrathecal administration, intranasal administration, intra-pulmonary administration, rectal administration, vaginal

administration, oral administration, and topical administration, and the like. All such routes may be suitable for administration of these compositions, and may be selected depending on the agent used, patient and condition treated, and similar factors by an attending physician. Routes of administration may be combined within the course of repeated therapy.

The composition containing the nucleic acid molecule can contain an optional agent that facilitates polynucleotide uptake in a cell, and a suitable pharmaceutically acceptable carrier. The specific formulation of the pharmaceutical composition depends upon the form of the active agent. Suitable pharmaceutically acceptable carriers facilitate administration of the nucleic acid molecule-containing compositions of this invention, but are physiologically inert and/or nonharmful. Carriers may be selected by one of skill in the art. Such carriers include but are not limited to, sterile saline, phosphate buffered saline, dextrose, sterilized water, glycerol, ethanol, lactose, sucrose, calcium phosphate, gelatin, dextran, agar, pectin, peanut oil, olive oil, sesame oil, and water and combinations thereof. Additionally, the carrier or diluent may include a time delay material, such as glycerol monostearate or glycerol distearate alone or with a wax. In addition, slow release polymer formulations can be used. The formulation should suit also the mode of administration. Selection of an appropriate carrier in accordance with the mode of administration is routinely performed by those skilled in the art.

Where the composition contains a nucleic acid molecule, e.g., a DNA molecule, plasmid, viral vector, or recombinant virus, or multiple copies of the nucleic acid molecule or different nucleic acid molecules, etc., as described above, the composition may desirably be formulated as a naked polynucleotide with only a carrier. Alternatively, such compositions desirably contain optional polynucleotide facilitating agents or co-agents, such as a local anaesthetic, a peptide, a lipid including cationic lipids, a liposome or lipidic particle, a polycation such as polylysine, a branched, three-dimensional polycation such as a dendrimer, a carbohydrate, a cationic amphiphile, a detergent, a benzylammonium surfactant, or another compound that facilitates polynucleotide transfer to cells. Non-exclusive examples of such facilitating agents or co-agents useful in this invention are described in U. S. Patent Nos. 5,593,972; 5,703,055; 5,739,118; 5,837,533 and International Patent Publication

No. WO96/10038, published April 4, 1996; and International Patent Publication No WO94/16737, published August 8, 1994, which are each incorporated herein by reference.

5 When the facilitating agent used is a local anesthetic, for example, bupivacaine, an amount of from about 0.1 weight percent to about 1.0 weight percent based on the total weight of the polynucleotide composition can be used. See, also, International Patent Publication No. WO98/48780 for delivery in vesicular complexes; and International Patent Publication No. WO99/21591, which teaches the incorporation of benzylammonium surfactants as co-agents, administered in an  
10 amount of between about 0.001 to about 0.03 weight %, the teachings of which are hereby incorporated by reference.

Still additional components that may be present in any of the compositions are preservatives, chemical stabilizers, or other antigenic proteins. Typically, stabilizers and preservatives are optimized to determine the best  
15 formulation for efficacy in the target human or animal. Suitable exemplary preservatives include chlorobutanol, potassium sorbate, sorbic acid, sulfur dioxide, propyl gallate, the parabens, ethyl vanillin, glycerin, phenol, and parachlorophenol. Suitable stabilizing ingredients which may be used include, for example, casamino acids, sucrose, gelatin, phenol red, N-Z amine, monopotassium diphosphate, lactose,  
20 lactalbumin hydrolysate, and dried milk.

In addition, other agents which may function as transfecting agents and/or replicating agents and/or inflammatory agents and which may be co-administered with the composition of this invention, include growth factors, cytokines and lymphokines such as alpha-interferon, gamma-interferon, platelet derived growth  
25 factor (PDGF), colony stimulating factors, such as G-CSF, GM-CSF, tumor necrosis factor (TNF), epidermal growth factor (EGF), and the interleukins, such as IL-1, IL-2, IL-4, IL-6, IL-8, IL-10 and IL-12. Further, fibroblast growth factor, surface active agents such as immune-stimulating complexes (ISCOMS), Freund's incomplete adjuvant, LPS analog including monophosphoryl Lipid A (MPL), muramyl peptides,  
30 quinone analogs and vesicular complexes such as squalene and hyaluronic acid may also be administered in conjunction with the compositions of the invention.



The pharmaceutical compositions may also contain other additives suitable for the selected mode of administration of the composition. Thus, these compositions can contain additives suitable for administration via any conventional route of administration, including those mentioned above. The composition of the invention may also involve lyophilized polynucleotides, which can be used with other pharmaceutically acceptable excipients for developing powder, liquid or suspension dosage forms, including those for intranasal or pulmonary applications. See, e.g., Remington: The Science and Practice of Pharmacy, Vol. 2, 19<sup>th</sup> edition (1995), e.g., Chapter 95 Aerosols, the teachings of which are hereby incorporated by reference. Routes of administration for these compositions may be combined, if desired, or adjusted.

The compositions of the present invention, when used as pharmaceutical compositions, can comprise about 1 ng to about 20 mgs of nucleic acid molecules e.g., plasmids, viral vectors, recombinant viruses, and mixtures thereof. In some embodiments, the compositions contain about 10 ng to about 10 mgs of nucleic acid molecule of the invention. In other embodiments, the pharmaceutical compositions contain about 0.1 to about 500  $\mu$ g nucleic acid molecule of the invention. In some embodiments, the compositions contain about 1 to about 350  $\mu$ g nucleic acid molecule of the invention. In still other embodiments, the pharmaceutical compositions contain about 25 to about 250  $\mu$ g of the nucleic acid molecule of the invention. In some embodiments, the compositions contain about 100  $\mu$ g of the nucleic acid molecule of the invention. Where the delivery agent is a live recombinant virus, a suitable vector-based composition contains between  $1 \times 10^2$  pfu to  $1 \times 10^{12}$  pfu per dose.

Similarly, the ligand is administered to the mammal to effect the silencing of the target gene *in vivo*. The ligand which is appropriate to the LBD in the chimeric fusion protein encoded by the nucleic acid molecule may be desirably administered by any of the above indicated routes of administration. However, to maintain the prolonged exposure of the nucleic acid molecule-transfected cells to the ligand, the ligand is administered in a conventional delayed releasing implant. The implant is desirably placed close in proximity to the target cells and is designed to release an effective amount of the ligand to the cells for a sufficient time to repress the

target gene expression. The effective amount and sufficient times for this use are expected to be similar to those stated above for *in vitro* use.

By following this method, it is possible to create a knock-out laboratory animal in which a target gene is stably suppressed. Such animals are  
5 useful in biological research. Also, by following this method, a gene therapy patient may have the undesirable expression of a target gene turned off.

If further manipulation is desired either in the knock-out animal or in a human patient, such as by reactivating the silenced gene, the reactivating molecule may be delivered *in vivo*. In yet a similar manner, the administration of the  
10 reactivating compound, its dosage, and the time period during which the mammal's cells are exposed to the reactivating compound that disrupts the binding of the complex at the target gene, effects DNA methylation of the target gene or effects histone acetylation will be determined by one of skill in the art.

The above dosage ranges for administering the nucleic acid molecules  
15 of this invention *in vivo* are guidelines only. The amount of the pharmaceutical composition in a dosage unit employed will be determined empirically, based on the response of cells *in vitro* and response of experimental animals to the compositions of this invention. It will be appreciated that optimum dose, time of exposure, route of administration, and need for readministration and formulations of nucleic acid  
20 molecule, ligand and/or optional reactivating compound for performance of this method to effect *in vivo* target gene silencing and reactivation may be determined by one of skill in the art, taking into account the target gene, the condition being treated, its severity, complicating conditions, and such factors as the age, and physical condition of the mammalian subject, the employment of other active compounds, and  
25 the like.

This invention provides a role for the KRAB-zinc finger protein (KRAB-ZFP) superfamily of transcriptional repressors in sequence-specific establishment of stable gene silencing. The KRAB-ZFPs selectively bind to cognate cis-regulatory elements and recruit the KAP-1 corepressor to the targeted locus.  
30 Because KAP-1 is obligatory for KRAB-mediated repression, the effector molecules of silencing are likely due to the network of proteins that interact with KAP-1. The examples below demonstrate that the KRAB-KAP1 repression system used in this

invention is a physiologically relevant targeting mechanism for HP1 proteins and, in addition, coordinates other activities which are likely to be required for silencing, i.e., the recruitment of the NuRD-HDAC complex (Schultz *et al.*, 2001, cited above) and recruitment of a novel histone H3-MeK9 methyltransferase, SETDB1.

5                   A comprehensive analysis of clonal NIH3T3 cell lines that contain these two plasmids are detailed in the following examples and support the following conclusions:

(1)       The KRAB domain is able to coordinate machinery for strong transcriptional repression of an integrated, target gene, e.g., a chromatinized RNA pol  
10       II transcribed target gene.

(2)       Repression is relatively short-range as a linked promoter ~2.8 kbp away is unaffected. The mechanism of KRAB-mediated repression functions over relatively short distances in chromatin. This property of a repression domain, i.e. the ability to mediate long-range versus short-range repression has emerged as a key  
15       determinant of the biological function for a repressor protein (Arnosti *et al.*, 1996 *EMBO J.*, 15:3659-3666; Gray and Levine, 1996 *Genes Dev.*, 10:700-710). Processes such as pattern formation, boundary determination, and control of cross-talk between closely linked genes are strongly dependent upon the distance over which a particular repressor will exert influence. The physical association of the KPHBD, KAP1, and  
20       HP1 proteins appears to only span a few nucleosomes, since the 3' end of the luciferase coding region, which is ~1.2 kbp away in the plasmids described in the examples, was devoid of any cross-linkable protein (KRAB/KAP1/HP1) in these assays.

(3)       Repression is accompanied by a highly localized chromatin  
25       compaction in the promoter region as judged by restriction endonuclease sensitivity, and spatial recruitment of the gene to a sub-nuclear region enriched in condensed heterochromatin. In this system, KRAB-mediated short-range repression was accompanied by highly localized chromatin compaction. The KAP1 corepressor and HP1 $\alpha/\gamma$  proteins are physically associated with the repressed gene in a highly  
30       localized manner as judged by ChIP assays. The inventors observed a physical association of the KPHBD protein at and around the PAX3 DNA recognition sequences. Furthermore, the KAP1 and HP1 proteins were enriched at DNA

sequences encompassing the nearby TK promoter region. The inventors also detected spatial relocation of the gene in the interphase nucleus to a block of constitutive condensed chromatin. Intuitively, such a long-range re-localization adjacent to a large block of A-T rich condensed chromatin would be expected to have long-range silencing effects on the gene. While one caveat to this result is the potential presence of a cryptic insulator/boundary element between the luciferase and zeocin genes, the inventors theorize that silencing mediated by a short-range repressor may occur by highly-localized looping into heterochromatic environments (Seum *et al.*, 2001, *EMBO J.*, 20:812-818). Such a scenario has been hypothesized to occur as a consequence of the formation of highly localized and repeating lattice units of chromatin that recognizes a complementary surface in constitutive heterochromatin (Singh and Huskisson, 1998 *Dev. Genet.*, 22:85-99).

(4) The induced, silent state of the transgenic reporter is mitotically heritable in the absence of hormone for at least 40 population doublings as judged by clonal analysis of the sub-populations. The most striking component of this invention is that KRAB-KAP1-HP1 mediated gene silencing transgene silencing is mitotically heritable in cell culture. This is the hallmark property that distinguishes HP1-dependent PEV from other mechanisms of repression. Instead of physical linkage (via translocation or de novo integration) of the euchromatic gene to adjacent heterochromatin as occurs in PEV, transient targeting of HP1 to a transcribed locus occurs using the KRAB-KAP1 system. The results of the examples below suggest that a pulse of KRAB-KAP-HP1 protein induces a stable, silenced state that can be detected at high frequency in clonal sub-populations following growth of more than 40 cell generations in the absence of ligand/inducer. This result was observed in three independent clonal lines derived from the original KPHBD-luciferase reporter transfection. Since it is highly likely that each cell line contains a unique luciferase transgene integration site, the observed effect is not due to a unique position effect in a single clone. Because, it is well established that the biological effects of HBD fusion proteins are readily reversible following removal of ligand, the stable silencing is being maintained in the absence of the KPHBD DNA binding activity. This is further supported by the observation that SNAG domain-mediated repression (which

does not involve HP1) is completely reversible by hormone withdrawal under the same experimental conditions.

5 The inventors first observed these apparently stably silenced sub-populations in mass culture over short time periods following 4-OHT treatment and withdrawal, suggesting the potential artifact of residual 4-OHT present in the culture medium. However, the single cell clonal analysis eliminates this concern as each clone is subjected to extensive washing, growth medium changes and trypsinization/ subculturing all over the course of ~30 days in the absence of 4-OHT. Analysis of the sub-clones derived from 4-OHT treatment showed a growth rate and zeocin resistance  
10 comparable to the parental clone. These data strongly argue that lack of luciferase activity is not due to selective growth inhibition, deletion, or any other genetic mechanisms of transgene silencing. This was confirmed by the finding that the luciferase gene could be strongly reactivated by transient expression of a PAX3-VP16 protein in the stably repressed clones. Moreover, the stably silenced clones can be  
15 strongly and synergistically reactivated (>100-fold in some cases), by 5-azacytidine + trichostatin-A (TSA) treatment when the drugs are administered in a specific sequence. These characteristics clearly show that the luciferase transgene is intact and functional in the stably silenced clones.

Moreover, these examples provide evidence that DNA methylation is  
20 likely responsible for the stably silenced state. Establishing the physical state of both the DNA and the chromatin at the luciferase locus in these stably silenced clones permits screening for agents that disrupt that state to reverse silencing. Two additional features have been noted. First, in three independent reporter clones, the duration of initial 4-OHT treatment (2, 4, or 6 days) directly correlated with the  
25 apparent frequency of stably repressed clones present in the population. One interpretation of this is that a passage through the cell cycle in the presence of the repressive signal enables and/or enhances the fixation of the stably silenced state. Alternatively, there may be a discrete cell cycle phase that is permissive for establishing the stably silenced state, thus repeated cycling through this window in the  
30 presence of the repressive signal would serve to sequentially enrich for the sub-population in the silenced state.

(6) Stable silencing is apparently specific for the KRAB-KAP1-HP1 mechanism. Other repression domain fusions (which do not bind HP1) do not induce a heritable, silenced state. In the examples, silencing did not appear to spread along the template into the adjacent zeocin transcription unit. Spreading is commonly observed in pericentromeric heterochromatin-mediated PEV in *Drosophila*. Since the inventors continually selected for zeocin expression (via drug selection) during growth of the single cell clones after the 4-OHT pulse, this selection pressure may have both impeded spreading and the establishment of a larger domain of silencing. Since under positive selection pressure, these clones would have been eliminated from the population, the clonal analysis was repeated in the absence of zeocin to determine if a higher frequency of stably silenced clones could be observed. However, the same frequency of stably silenced clones was observed in the absence of zeocin when compared to the +zeocin experiment. This provides additional evidence that silencing of euchromatic genes via localized recruitment of HP1 may be unique to KRAB repressor domain and also that KRAB is fundamentally different from the other repressor domains.

Together, these results strongly suggest a model for HP1-dependent silencing and variegation of a euchromatic gene in a mammalian cell line.

The examples further demonstrate that the interaction between KAP-1 and SETDB1, a novel histone H3, Lysine-9 specific histone methyltransferase contributes to this coordinated repression mechanism. Most interesting, the ChIP experiments of silent and active clones revealed the enrichment of DNA spanning the TK promoter and transcription initiation site in KAP-1, SETDB1, HP1 and H3-MeK9 immunoprecipitates from cell clones containing a stably silenced luciferase transgene. Thus, the KRAB-KAP-1 repression system is one of the best characterized mammalian systems of gene-specific silencing at euchromatic genes by targeting HP1 proteins.

The KRAB-KAP-1 repression complexes coordinate biochemical activities that induce localized assembly of higher order chromatin structure to repress transcription. Increased DNA methylation in the stably silenced clone also centered around the proximal promoter region bound by these proteins. That this DNA methylation contributes to the silencing is suggested by the potent synergistic

reactivation of the silent locus, i.e., silenced luciferase transgenes, by the combined and specific administration of 5-azacytidine +TSA. This result signifies that stable silencing of gene expression by KRAB-ZFPs may ultimately result in DNA methylation of key cis-regulatory elements, i.e., H3-Mek9 methylation, and HP1 recruitment on a euchromatic gene.

The role of histone methylation in these processes has been revolutionized by the discovery that proteins with the highly conserved SET domain function as lysine-specific histone methyltransferases. In the experiments described herein, the inventors provide evidence that SETDB1, a novel SET domain protein, is a highly selective histone H3 lysine-9 methyltransferase. Moreover, the inventors demonstrate that SETDB1 can be targeted to a stably silenced euchromatic locus via the KAP-1 corepressor of the KRAB-ZFP family of sequence-specific transcriptional repressors.

DNA methylation likely plays an active role in maintaining the histone-directed machinery at the silent locus. This could be accomplished in two ways. First, the MBD2/3 component of the NuRD histone deacetylase complex directly could directly bind the methylated DNA and maintain HDAC activity at the locus. Second, in addition to the catalytic SET domain, the SETDB1 protein also encodes a CpG DNA Methyl Binding Domain which, if functional, would maintain the H3-MeK9 activity at the locus (Schultz *et al.*, 2002, cited above). That this may occur is supported by our preliminary observation that 5AZA+TSA reactivation is accompanied by rapid loss of SETDB1 and H3-MeK9 at the locus as assessed by ChIP assays (data not shown).

A KRAB-ZFP binds to its cognate recognition sequence and then recruits KAP-1 to form a scaffold that coordinates the assembly histone deacetylases, histone methylases, and the deposition of HP1 proteins to silence gene expression by either (i) creating a localized heterochromatic environment in nuclear domains cytologically consistent with euchromatin, or (ii) relocating the gene to an established nuclear heterochromatin territory via some looping or tethering mechanism. In either event, KAP-1 coordinates biochemical activities that induce the assembly of higher order chromatin structure.

Until recently, most insights into HP1 protein function and the role of heterochromatin in gene regulation have come from the study of position effect variegation (PEV) in *Drosophila*. As in PEV, large blocks of constitutive heterochromatin are also highly enriched in the HP1 protein and its anchor, histone H3-MeK9, which together function as exquisite dose-sensitive silencers of an adjacent euchromatic transgene. Once established, this silencing is mitotically heritable through many cell divisions giving rise to the classic variegated gene expression phenotype in the adult organs.

## EXAMPLES

The following examples are provided to illustrate the production and activity of representative compounds of the invention and to illustrate their performance in a screening assay. One skilled in the art will appreciate that although specific reagents and conditions are outlined in the following examples, these reagents and conditions are not a limitation on the present invention.

### EXAMPLE 1: HORMONE REGULATABLE CHIMERIC REPRESSOR PROTEINS

The inventors devised a two-plasmid system composed of a regulatable chimeric repressor, and a synthetic reporter gene, which is recognized by the DNA binding domain of the chimeric repressor. As detailed below, this two-plasmid system comprises a valid, hormone and DNA binding dependent repression model.

#### A. Repressor Plasmids

The pcKRAB-PAX3 (pcKP) plasmid was constructed by fusing the 90 amino acid KRAB domain from the Kox1 gene (Margolin *et al.*, 1994 *Proc. Natl. Acad. Sci. USA*, 91:4509-4513) to the NH<sub>2</sub> terminus of the PAX3 DNA binding domain (DBD) (Ayyanathan *et al.*, 2000, cited above). The KRAB domain binds KAP1 and is a very strong, DNA-binding dependent repressor *in vivo*. The PAX3 DBD binds DNA as a monomer, recognizes an extended non-degenerate DNA binding site, is easily detectable using PAX3 antibodies, and is neutral when bound to DNA in the absence of a fused effector domain (Fredericks *et al.*, 2000 cited above;



Fredericks *et al.*, 2001 cited above; Fredericks *et al.*, 1995 *Mol. Cell. Biol.*, 15:1522-1535).

To make the pKB chimeric repressor hormone regulatable, the pcKRAB-PAX3-HBD (KPHBD) plasmid was constructed by fusing the tamoxifen mutant hormone binding domain of the murine estrogen receptor (ERHBD<sup>TM</sup>) in frame with the COOH-terminus of the PAX3 DNA binding domain (DBD) at a unique EcoR I site, present just upstream of the stop codon in the pcKP plasmid. The ERHBD<sup>TM</sup> domain contains a single amino acid substitution (G<sub>525</sub>R) which renders it 1000-fold less responsive (essentially unresponsive) to serum estrogens, and contains no intrinsic transcriptional activation potential (Littlewood *et al.*, 1995 *Nucl. Acids Res.*, 23:1686-1690). See Fig. 1A.

The pcKRAB(DV)-PAX3-HBD (K(DV)PHBD) plasmid was constructed by sub-cloning a Hind III/BamH I fragment containing the mutant KRAB (DV<sub>18,19</sub>AA) domain (Margolin *et al.*, 1994, cited above) into pcKRAB-PAX3-HBD plasmid, replacing the wild type KRAB domain. The mutant KRAB (DV<sub>18,19</sub>AA) lacks repression activity and fails to bind KAP1.

These chimeras can be stably expressed in the cell, are constitutively nuclear localized, but are inactive for DNA binding in the absence of hormone presumably due to steric hindrance or masking of the adjacent DBD by the ERHBD<sup>TM</sup>.

For comparison purposes, artificial repressors of identical architecture were constructed using the well-characterized modular repression domains (RD) from WT1 (Madden *et al.*, 1991 *Science*, 253:1550-1553), BTB/POZ from the PLZF gene (Li *et al.*, 1997 *J. Biol. Chem.*, 272:27324-27329), Engrailed from the GFI protooncogene (Jaynes and O'Farrell, 1991, *EMBO J.*, 10: 1427-1433) and SNAG from GFI-1 protooncogene (Ayyanathan *et al.*, 2000, cited above; Zweidler-Mckay *et al.*, 1996 *Mol. Cell. Biol.*, 16:4024-4034). All PCR derived nucleotide sequences and the appropriate fusion junctions were confirmed by sequencing both DNA strands. The resulting plasmids were pcSNAG-PAX3-HBD (SPHBD), pcPLZF(POZ)-PAX3-HBD (PPHBD), pcEngrailed-PAX3-HBD (EPHBD) and pcWT1-PAX3-HBD (WPHBD).

Each RD-PAX3-HBD fusion gene was expressed from a CMV promoter vector, which also contained a neomycin resistance cassette (Neomycin<sup>R</sup>) that enabled selection of stably transfected cells. Each chimeric protein was stably expressed in cells as shown by COS-1 transfection and immunoprecipitation using antibodies directed against PAX3 (Fredericks et al., 1995, cited above). Briefly, for each engineered plasmid, the stable expression of the chimeric protein was monitored by immunoprecipitation of (<sup>35</sup>S)-L-methionine-labeled whole cell extracts from transiently transfected COS-1 cells with  $\alpha$ -PAX3 IgG (gel not shown). In addition, each protein was localized to the nucleus and displayed high affinity for the PAX3 recognition sequence as shown by gel shift assays using transfected cell extracts (data not shown).

#### B. Reporter Plasmid

The CD19-TK-LUC plasmid, a derivative of pSP64, contained six repeats of the high affinity PAX3 DNA binding motif derived from the CD19 gene, followed by a herpes simplex virus (HSV) minimal thymidine kinase (TK) promoter that controls the basal expression levels of the luciferase gene (Schafer et al., 1994 *Nucl. Acids Res.*, 22:4574-4582). A Pvu II fragment (zeocin<sup>R</sup> cassette) from pcDNA3.1/Zeo plasmid (Invitrogen) containing 1275 to 3024 bps, which encompasses the SV40 and EM-7 promoters, zeocin<sup>R</sup> gene and SV40 polyadenylation sequence, was cloned into a unique Pvu II site in the CD19-TK-LUC plasmid. This gene insert provided both a selection marker allowing generation of stable cell clones, and a second transcription unit, linked to the luciferase gene. Transformants were selected on LB plates containing 25  $\mu$ g/ml of zeocin (Invitrogen). Restriction mapping of the resulting CD19-TK LUC-Zeo<sup>R</sup> plasmid indicated that the zeocin<sup>R</sup> cassette was present in a tail to tail orientation with respect to the luciferase gene (Fig. 2A). Since the promoters that drive these two genes are located about 2.8 kbp apart, they can be used to test the distance requirement for KRAB-KAP1 mediated repression *in vivo*.

The designation of the CD19-TK-LUC gene as euchromatic is based upon the following functional criteria: 1) the stably transfected clones were selected both for, and display a high basal level of luciferase activity, 2) all clones both express, and show physical linkage to the zeocin resistant cassette, 3) the reporter

genes are packaged into a regular, nuclease accessible chromatin structure, and 4) interphase FISH shows the reporter transgene to be present predominantly in euchromatic chromosome territories in the nucleus. Thus, while designation of a gene (or a region) as euchromatic is historically based upon cytological analyses, clearly, in the clones selected for study, the reporter plasmids have integrated into a region permissive for a high level of expression.

*C. 4-OHT-Dependent Repression Of A Transient Reporter*

To determine if these chimeric proteins behaved as hormone dependent repressors and determine the 4-OHT-dependent repression potentials of the chimeric repressor proteins,  $2 \times 10^5$  NIH3T3 cells in 60 mm dishes were transiently transfected with 1  $\mu$ g of each expression plasmid, 0.5  $\mu$ g of CD19-TK-LUC or CD19-TK-LUC-Zeo<sup>R</sup> reporter plasmid, and 0.25  $\mu$ g of pCMV-LacZ plasmids, using lipofectAMINE (Life Technologies). Six hours post-transfection, duplicate dishes of cells were treated with fresh growth medium (DMEM plus 10% calf serum) containing either 0.1% ethanol (-OHT dishes) or 500 nM 4-OHT (Research Biochemicals International, Natick, MA) (+OHT dishes) for 24 hours. Whole cell extracts (cell lysates) were assayed for luciferase activities and normalized to  $\beta$ -galactosidase values for transfection efficiency. Fold repression was determined as the ratio of normalized light units in -OHT treated cells to that in +OHT treated cells.

For analysis of the recovery of luciferase expression after 4-OHT removal, cells were first seeded at  $0.5 \times 10^4$  in 60-mm dishes. Duplicate dishes of cells were then treated with either 500 nM 4-OHT (+OHT) or 0.1% ethanol (-OHT) for the indicated time. At the end of the induction period, one dish from each of the -OHT and +OHT treated cells was harvested while the remaining dishes were subjected to three washes in a 24-hr period each day. A wash was defined as two changes of 4 ml of DMEM followed by addition of 4 ml of complete growth medium. At the end of 4th day, the cells were harvested by trypsinization and re-plated into new dishes. Cell washings were continued for an additional 4 days. Cells were harvested daily and normalized luciferase activities determined.

The results are shown in Fig. 1B. The CD19-TK-LUC-Zeo<sup>R</sup> plasmid showed a high basal level of luciferase activity ( $>190,000$  light units/ $\mu$ g plasmid in 24 hours) that was unaffected by co-transfection with the vector pcDNA3. However, 4-

OHT dependent repression of the luciferase reporter by each RD-PAX3-HBD plasmid was observed. The KPHBD protein showed maximal repression (>10 fold), while the SPHBD, EPHBD, PPHBD and WPHBD expression constructs elicited more moderate levels of repression (~ 3 to 6 fold). As expected, the K(DV)PHBD protein was almost completely inactive for repression activity. None of the RD-PAX3-HBD proteins showed repression activity on a TK-LUC-Zeo<sup>R</sup> reporter, which lacked PAX3 binding sites (data not shown). Importantly, the RD-PAX3-HBD system was tightly regulatable by 4-OHT. Even for the most powerful repressor (KPHBD), very little change in basal luciferase activity was observed in the absence of hormone treatment (Fig. 1B). Thus, this two-plasmid system comprises a valid, hormone- and DNA-binding dependent repression model.

#### EXAMPLE 2: THE KRAB-PAX3-HBD PROTEIN UNIQUELY ASSOCIATES WITH HETEROCHROMATIN PROTEINS THROUGH THE KAP1 COREPRESSOR

To determine if the KRAB-PAX3-HBD fusion protein or any of the other chimeric repressors would bind to a KAP1-HP1 complex, the RD-PAX3-HBD protein expression was confirmed by transient transfection of COS-1 cells with the expression plasmids encoding the chimeric PAX3 repressor plasmids followed by immunoprecipitation of the (<sup>35</sup>S)-L-methionine labeled whole cell extracts with  $\alpha$ -PAX3 IgG (Ryan *et al.*, 1999 *Mol. Cell. Biol.*, 19:4366-4378). Whole cell lysates from COS-1 cells transfected with either the KPHBD, K(DV)PHBD or SPHBD expression plasmids were incubated with 5  $\mu$ g of either recombinant GST, GST-HP1 $\alpha$  or GST-HP1 $\gamma$  proteins immobilized on a GSH-sepharose affinity chromatography column. The HP1-associated proteins retained by the GST-HP1 resins were eluted, fractionated in 12% SDS-PAGE, and were analyzed by Western immunoblotting sequentially using  $\alpha$ -PAX3 and  $\alpha$ -KAP1 IgG specific antibodies (Gels not shown).

As seen in the graphics of Fig. 1C, the KPHBD protein interacts with the KAP1 corepressor through a direct interaction between the KRAB box and the RBCC domain of KAP1 (Peng *et al.*, 2000 *J. Mol. Biol.*, 295:1139-1162). The KAP1 corepressor in turn interacts with the chromoshadow (CSD) domain of the HP1 protein family through a core PxVxL motif (Lechner *et al.*, 2000 *Mol. Cell. Biol.*,

20:6449-6465; Ryan *et al.*, 1999 cited above). Both GST-HP1 proteins efficiently retained KAP1 and the wild-type KRAB-PAX3-HBD protein derived from cell extracts. This interaction was specific, as the mutant KRAB domain (DV<sub>18,19</sub>AA), the SNAG-PAX3-HBD fusion, or any of the other fusion proteins (data not shown) failed to bind endogenous KAP1. Thus, the KRAB-PAX3-HBD protein retains the ability to efficiently bind a KAP1-HP1 complex. KAP1 may be constitutively bound to the fusion protein *in vivo*. The other repression domains do not bind either KAP1 or any other endogenous complexes that have the capability of interacting with HP1 proteins in this assay. Thus, this set of chimeric repressors allows a comparison of HP1-mediated and HP1-independent mechanisms of gene silencing.

### EXAMPLE 3: 4-OHT-DEPENDENT REPRESSION OF CHROMATINIZED REPORTER TRANSGENES

The strategy for creating mammalian cell lines with integrated luciferase reporter transgenes is depicted in Fig. 2A. The DNAs for the CD19-TK-LUC-Zeo<sup>R</sup> luciferase reporter and each of the RD-PAX3-HBD fusions were co-transfected into murine NIH3T3 fibroblasts. Clonal populations of NIH3T3 cell lines that contain only a stably integrated CD19-TK-LUC-Zeo<sup>R</sup> reporter plasmid were generated by selection in growth medium containing 100 µg/ml of zeocin. NIH3T3 cell lines containing a stably expressed chimeric repressor and the CD19-TK-LUC-Zeo<sup>R</sup> reporter plasmid were generated by co-transfection and selection in growth medium containing 500 µg/ml G418 and 100 µg/ml zeocin. Following G418+zeocin selection, clones were isolated from independent dishes at the end of two weeks via ring cloning, and expanded into mass population.

The basal luciferase activities and the 4-OHT-dependent repression of the integrated luciferase reporter gene were measured for expanded populations of clonal cells. The luciferase activities were normalized to protein concentration as determined by Bradford assay and expressed as light units/O.D. at A<sub>595</sub>. The clones with the highest normalized luciferase activity, suggesting that the luciferase transgene had integrated at a locus permissive for transcription, were then tested for 4-OHT dependent repression activity. Stable expression of the chimeric RD-PAX3-

HBD proteins was tested by immunoprecipitation of the ( $^{35}$ S)-L-methionine labeled whole cell extracts with  $\alpha$ -PAX3 IgG as described above. As controls, stable cell lines were generated that contain only the chromatin-integrated luciferase reporter (designated as "CL" clones).

5           At least 48 independent clonal cell lines both for CL and for each RD-PAX3-HBD fusion were tested for both basal luciferase activity and 4-OHT-dependent repression. A representative selection of five clones each for CL, KPHBD, K(DV)PHBD and SPHBD is shown in Fig. 2B and 2C. The normalized luciferase activities, among clones, varied from  $\sim 10^3$  to  $10^6$  (light units/1 O.D. at  $A_{595}$  protein) (Fig. 2B). This variation likely reflects the position effects of integration sites observed with different chromosomal loci.

10           As expected, the CL clones, which do not contain a stable RD-PAX3-HBD gene show no response to 4-OHT. However, strong 4-OHT dependent repression was observed in a good fraction of the KPHBD transfected clones: KPHBD21 cell line showed strong repression ( $\sim 16$ -fold), while the KPHBD 29 clone showed a very low repression activity ( $\sim 2$ -fold). The KPHBD 30, KPHBD 33 and KPHBD 36 clones manifested moderate repression activity ( $\sim 3$  to 5-fold).

15           In contrast, each of the K(DV)PHBD transfected clones possessed a high basal luciferase activity that was unaffected by 4-OHT treatment. However, the SPHBD transfection produced clones that showed 6 to 7-fold repression in response to 4-OHT (Fig. 2C). Each clone that demonstrated 4-OHT responsive changes in luciferase activity was tested for stable expression of the RD-PAX3-HBD chimeric repressor fusion protein by metabolic labeling with ( $^{35}$ S)-L-methionine and immunoprecipitation with  $\alpha$ -PAX3 IgG. The expression of the RD-PAX3-HBD proteins was observed in KPHBD, K(DV)PHBD, and SPHBD stable cell lines that display 4-OHT dependent repression of luciferase activity. Expression of appropriate full-length protein was determined by corresponding *in vitro*-translated product. Each clone that demonstrated 4-OHT responsive changes in luciferase activity stably expressed the appropriate sized full-length RD-PAX3-HBD protein (data not shown).  
20  
25  
30           As expected, a range of expression levels was observed among the clones, which correlated roughly with the 4-OHT dependent repression potential of these cells.

In a similar manner, the stable cell lines were generated containing integrated luciferase reporter and expressing either a engrailed-PAX3-HBD (EPHBD), PLZF-POZ-PAX3-HBD (PPHBD) or WT1-PAX3-HBD (WPHBD) chimeric repressor protein. These cell lines also showed demonstrable levels of 4-OHT dependent repression.

From these observations, it can be concluded that: 1) 4-OHT treatment does not affect the basal expression of the reporter luciferase transgene when stably integrated at many different sites in the genome. 2) The KPHBD fusion protein is a powerful, hormone-dependent repressor of the integrated reporter transgene whereas a mutant version, which cannot bind KAP1 fails to demonstrate any appreciable levels of repression; and 3) NIH3T3 cells contain the machinery required to support SNAG domain mediated repression.

#### EXAMPLE 4: MOLECULAR CHARACTERIZATION OF A KRAB-PAX3-HBD STABLE CELL LINE

A battery of molecular genetic characterizations were performed on the KPHBD21 cell line.

(a) First, PCR using genomic DNA showed that the PAX3 binding sites were contiguous with the TK promoter (data not shown). A similar analysis showed that the SV40 promoter-zeocin<sup>R</sup> gene (zeocin<sup>R</sup> cassette) was also physically linked to the luciferase gene (data not shown).

(b) Second, quantitative Southern blotting suggested that an estimated 2 to 5 copies of the CD19-TK-LUC Zeo<sup>R</sup> were present (data not shown).

(c) Third, micrococcal nuclease digestion and Southern blotting showed a standard nucleosomal pattern using a probe for the TK promoter region. Moreover, this pattern was altered in response to 4-OHT suggesting a change in nucleosome positioning accompanied repression (data not shown).

(d) Fourth, the repression observed was strongly time- and 4-OHT concentration-dependent as shown in Fig. 2D.

(e) Quantitative Reverse Transcription Polymerase Chain Reaction (Q-RT-PCR) was performed as follows: KPHBD21 cells were treated with either 0.1% ethanol (-OHT) or 500 nM 4-OHT (+OHT) for 48 hours. Total RNAs were isolated

using TRIzol Reagent (Life Technologies) and oligo-dT primed first strand cDNAs made. Luciferase, neomycin<sup>R</sup>, and zeocin<sup>R</sup> mRNAs were amplified by PCR for the number of cycles indicated. Each 50 µl reaction contained 5 µl of either -OHT or +OHT first strand cDNAs, 1x PCR buffer, 10% DMSO, 250 µM dNTPs, 5 units of Taq DNA and 100 pmoles of respective primer-pairs (SEQ ID NOS: 1-6, respectively):

LUC2: 5' CAAGGATATGGGCTCAC 3' and

LUC3: 5' GACCTTTCGGTACTTCG 3';

NEO1: 5' TCAGCGCAGGGGCGCCCGGTTCTTT 3' and

NEO2: 5' ATCGACAAGACCGGCTTCCATCCGA 3';

ZEO1: 5' ATGGCCAAGTTGACCAG 3' and

ZEO2: 5' TCAGTCCTGCTCCTCG 3'.

Equal volumes of the products were analyzed on 1.5% agarose gels. The DNA fragments were Southern-blotted, pre-hybridized, and hybridized overnight with respective probes (Ayyanathan et al., 2000, cited above). Hybridized membranes were washed in 0.2xSSC/0.2% SDS at 50°C for 30 minutes and exposed to Kodak MR X-ray film for 4 hours. The signals were also quantified using ImageQuaNT® system (Molecular Dynamics).

4-OHT dependent repression occurred at the level of transcription as shown by decreases in the abundance of luciferase mRNA measured by quantitative RT-PCR (data not shown). Moreover, transcription at both the linked zeocin<sup>R</sup> locus, and the unlinked neomycin<sup>R</sup> locus were unaffected by 4-OHT treatment in the quantitative RT-PCR assays (Fig. 2E). Thus, the KPHBD21 cell line is a valid model for studying mechanisms of KRAB domain mediated transcriptional repression of an integrated gene that is assembled into chromatin structure. Moreover, KRAB mediated repression in this system appears to be highly localized, as a linked transcription unit (zeocin<sup>R</sup> cassette) ~2.8 kbp away from the repressor binding site is unaffected.

#### EXAMPLE 5: KRAB-PAX3-HBD PROTEIN INDUCES A HIGHLY LOCALIZED COMPACT CHROMATIN STRUCTURE



To begin to dissect the changes in chromatin structure that accompany repression in the KPHBD21 cell line, the inventors performed nuclease accessibility assays using known restriction enzyme sites in the CD19-TK-LUC Zeo<sup>R</sup> plasmid.

The *in vivo* analysis of chromatin structure was performed as follows:

5 KPHBD21 cells were treated with either 500 nM of 4-OHT (+OHT dishes) or 0.1% ethanol (-OHT dishes) for 4 days, and intact nuclei were prepared essentially as described (Mymryk *et al.*, 1997 *Methods*, 12:105-114). The nuclei pellets were resuspended in 10 mM Tris-HCl pH 7.4, 15 mM NaCl, 60 mM KCl, 0.15 mM Spermine, 0.5 mM Spermidine, 5 mM MgCl<sub>2</sub> and 5% (v/v) glycerol.

10 For the restriction enzyme accessibility assays, the nuclei from mock or 4-OHT treated cells were resuspended in appropriate 1x buffers in a 500 µl reaction volume and digested with 250 units of restriction enzymes BamHI, BglII, EcoRI, HindIII, SmaI and XhoI for 10 minutes at room temperature (RT). The endonuclease reaction was terminated by adding 100 µg of proteinase-K in 10 mM Tris-HCl pH 8.0;  
15 10 mM EDTA; 10 mM NaCl, and incubated at 37°C for 14-16 hours. These samples were phenol:chloroform extracted, and ethanol precipitated.

To ensure that only a limited digest of the nuclei occurred, resuspended DNA was cleaved to completion with BamHI, which created an internal control for each reaction. After purifying the DNA as above, the extent of cleavage at each site was  
20 quantitated via an indirect end-labeling technique using a luciferase gene primer. Equal amounts of the DNA samples were taken for reiterative, primer extension PCR reactions with the following <sup>32</sup>P-labeled primers:

LUC1: 5' TCCAGGAACCAGGGCGTATCTCT 3' (SEQ ID NO: 7) or  
ZEO2 (SEQ ID NO: 6), or NEO2 (SEQ ID NO: 4).

25 Primer extension products were extracted with phenol:chloroform and ethanol precipitated. Dried DNA pellets were dissolved in formamide gel loading buffer, electrophoresed in 7M Urea/5% acrylamide gels in 0.5x TBE buffer along with radiolabeled ΦX174-Hae III size marker. The gels were fixed in 10% acetic acid, dried and autoradiographed (data not shown).

30 Endonuclease accessibility displayed a differential at the luciferase locus. Restriction endonuclease sensitivity at the HSV TK promoter controls basal expression of luciferase gene. Nuclei isolated from -OHT (0.1% ethanol) or +OHT

(500 nM of 4-OHT) treated KPHBD21 cells (for 4 days continuously) were digested with restriction endonucleases BamHI, BglII, EcoRI, HindIII, Sma I and XhoI. Subsequent to complete digestion of the isolated DNA with BamHI, reiterative primer-extension PCR reaction was done using radiolabeled LUC1 primer. The denatured products were resolved in a 7M Urea-5%PAGE and autoradiographed (not shown).

Each enzyme showed dramatic inhibition of cleavage in the 4-OHT treated nuclei. Thus, a region of 257 bp flanking the TK promoter and transcription initiation site in the luciferase gene is converted to a compact, nuclease-resistant structure upon 4-OHT treatment.

As a control, similar experiments for restriction endonuclease sensitivity (NcoI, StuI and XmnI) were performed on sites in the SV40 promoter, which controls basal expression of the unlinked neomycin<sup>R</sup> gene determined by using radiolabeled NEO2 primer using molecular weight marker (ΦX174 Hae III digest) (data not shown). Strong cleavage was observed at both the Nco I and Stu I sites in the presence of 4-OHT.

Similar sites (NcoI, StuI and XmnI) in the SV40 promoter, which controls basal expression of the zeocin<sup>R</sup> gene were also assayed (data not shown) using radiolabeled ZEO2 primers. These sites also showed equal accessibility regardless of 4-OHT treatment. Thus, the compact chromatin structure induced by KPHBD binding to the PAX3 sites near the TK promoter is not established at a promoter 2.8 kbp apart. This suggests that KRAB-mediated repression is highly localized in the context of a chromatinized reporter.

#### EXAMPLE 6: RECRUITMENT OF KAP1 AND HP1 TO THE REPORTER LUCIFERASE TRANSGENE

To determine if known components of the KRAB-KAP1 silencing machinery are bound to the target gene, comprehensive chromatin immunoprecipitation (ChIP) analyses of the locus was performed. The KRAB box dependent recruitment of KAP1 and HP1 proteins to the integrated luciferase transgene was shown by direct quantitation of scanned gel images (data not shown).

Chromatin associated proteins were chemically cross-linked to DNA *in vivo* with formaldehyde in mock or 4-OHT treated KPHBD21 and CL2 cells as follows: Cells were plated at  $5 \times 10^5$  cells/150-mm dish, treated continuously with either 500 nM 4-OHT (+OHT dishes) or 0.1% ethanol (-OHT dishes) for 4 days. DNA-protein complexes were chemically cross-linked and fixed *in vivo* with 1% formaldehyde (EM Biosciences) for 20 minutes at 37°C. PBS washed cells were resuspended in buffer I (10 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.5 mM EGTA, 0.25% Triton-X-100, 1 mM PMSF) and incubated at RT for 15 minutes. Cell lysates were clarified at 400 x g for 5 minutes. The pellet was resuspended in buffer II (10 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.5 mM EGTA, 200 mM NaCl, 1 mM PMSF), incubated at room temperature (RT) for 15 minutes, and centrifuged at 2000 x g for 5 minutes. The soluble chromatin was resuspended in buffer III (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 1 mM PMSF, 5 µg/ml of leupeptin, pepstatin and aprotinin) containing 1/3<sup>rd</sup> volume of glass beads (Biospec Products, Inc.).

Chromatin was sonicated on ice to average 400-600 bp fragments. The sonicated chromatin was adjusted to IP buffer conditions (20 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 200 mM NaCl, 0.5% Triton-X-100, 0.05% deoxycholate, 0.1% NP-40, 1 mM PMSF, 5 µg/ml of leupeptin, pepstatin and aprotinin) and clarified at 14,000 x g. The supernatant was pre-cleared with DNA-protein-A sepharose beads for 2 hours.

Soluble, sonicated chromatin fractions were immunoprecipitated (400 µl/IP) with 10 µg of α-PAX3 (Fredericks et al, 1995), affinity purified α-KAP1 raised against amino acids 20-418 (RBCC) (Schultz *et al*, 2001, cited above), α-HP1α and α-HP1γ IgG antibodies specific to the PAX3, KAP1, HP1α and HP1γ proteins. The HP1α and HP1γ reagents were monoclonal antibodies that were produced using purified, 6 HIS-tagged, full-length human antigens. These reagents do not cross react with the other human or mouse HP1 orthologues. Ten-percent volume of the clarified chromatin was saved as input. The immune complexes were recovered using fresh protein-A sepharose beads and washed as follows: twice with an ice-cold buffer protease inhibitor (20 mM Tris-HCl pH 8.0, 200 mM NaCl, 0.5% Triton-X-100, 0.05% deoxycholate, 0.1% NP40, 1 mM PMSF, 5 µg/ml of leupeptin, pepstatin and aprotinin) (buffer A) at 4°C, twice with buffer A containing 500 mM NaCl at RT,

twice with buffer A containing 250 mM LiCl at RT, and twice with TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) at RT. The immune complexes were stripped from the agarose beads and the protein-DNA cross links reversed by incubating for 6 hrs at 65°C in 50 mM Tris-HCl pH 8.0, 5 mM EDTA and 1% SDS. These samples were digested with proteinase-K, extracted with phenol:chloroform, and ethanol precipitated with 20 µg of glycogen as carrier.

Quantitative PCR reactions were carried out for 1% of the input DNA and immunoprecipitated DNAs using specific primer pairs that amplify various regions of the CD19-TK-LUC Zeo<sup>R</sup> locus (Fig. 2A). The following oligonucleotides: LUC1, LUC2, LUC3, ZEO1, ZEO2, NEO1, NEO2 (SEQ ID NOS: 7, 1, 2, 5, 6, 3, 4, respectively) and the following SEQ ID NOS: 8-16 respectively:

PBS5: 5' AGCGGGTGGTGGCGGGTGTC 3';

PBS3: 5' GAATACACGGAATTGGATCCG 3';

PBS1: 5' GATCGATAATTCGAGCTACTG 3';

PBS2: 5' GAGCTCGGTACCCGGGTCG 3';

PBS4: 5' GCCAATGACAAGACGCTGGG 3';

TKP1: 5' GCGCGGTCCCAGGTCCACTT 3';

SVP1: 5' CCAGTTCCGCCCATCTCTCCC 3';

CMP1: 5' ACGGGGATTTCCTCAAGTCTCC 3';

KRAB1: 5' TCCTCCCTGGTGAAGTCCAC 3'

were used in ChIP PCR reactions.

The SV40 promoters of the neomycin<sup>R</sup> and zeocin<sup>R</sup> cassettes, the CMV promoter that drives KPHBD expression, and downstream coding regions of the luciferase and zeocin<sup>R</sup> genes were used as specificity controls in the ChIP analyses.

The PCR reactions were electrophoresed in 1.5% agarose gels, stained with ethidium bromide, and photographed. PCR products for -OHT and +OHT treated cells were detected from ChIPs with α-PAX3, α-KAP1, α-HP1α and α-HP1γ IgG. The bands were quantified for fold enrichment using IQMac v1.2 software analysis of gel scans. The gels were Southern-blotted, hybridized with respective probes, autoradiographed, and also quantified by phosphorimager analysis.

The results are summarized in the following Table I and Fig. 3. No signal is reported as NS. The Row #s in the table below are shown beneath the relative positions of the schematic transgene fragments amplified in Fig. 3.

Table I

Cell Line	Primer Pairs	Base Pair	Fold enrichment by Antibodies				Row# for Fig. 3
			$\alpha$ -PAX3	A-KAP1	$\alpha$ -HP1 $\alpha$	$\alpha$ -HP1 $\gamma$	
KPHBD21	PBS3&PBS2	412	5.29	1.42	1.62	1.74	1
CL2	PBS3&PBS2	412	NS	NS	NS	NS	2
KPHBD21	PBS1&PBS2	344	10.64	10.99	4.56	3.44	3
KPHBD21	PBS3&PBS4	450	5.13	2.23	2.00	1.82	4
CL2	PBS3&PBS4	450	NS	NS	NS	NS	5
KPHBD21	TKP1&LUC1	257	2.78	9.55	27.00	1.47	6
KPHBD21	LUC2&LUC3	564	NS	NS	NS	NS	7
KPHBD21	ZEO1&ZEO2	374	0.94	0.65	1.01	0.93	8
KPHBD21	SVP1&ZEO2	618	NS	NS	NS	NS	9
KPHBD21	CMP1&KRAB1	314	NS	NS	NS	NS	10
KPHBD21	SVP1& NEO1	658	NS	NS	NS	NS	11

5

Similar results were observed in the phosphorimager analysis of the corresponding Southern blots (data not shown). Fragments corresponding to the three primer-pairs, which variably bracket the PAX3 binding sites, were found to be strongly enriched (5 to 10-fold) in the PAX3 immunoprecipitates (IPs) after 4-OHT treatment. This is good evidence that the HBD regulates KPHBD by limiting its DNA binding activity.

10

In addition to the DNA binding component, other components of the KRAB repression complex (i.e. KAP1, HP1 $\alpha$ , HP1 $\gamma$ ) were inducibly recruited to the target gene. This enrichment was most evident for the primer pair that directly flanks the PAX3 binding sites (PBS1 & PBS2). This fragment is enriched 10-fold, 11-fold and ~5-fold in PAX3, KAP1 and HP1 $\alpha$  IPs, respectively. The specificity of targeting by KPHBD is reflected in the fact that no signal (NS) was detected for any of these fragments in IPs of chromatin prepared from the CL2 reporter cell line, which lacks a

15

KPHBD transgene. The 257 bp fragment spanning the TK promoter and transcription initiation site is most highly enriched in KAP1 and HP1 $\alpha$  IPs, 10-fold and 27-fold respectively. This is dramatically different from HP1 $\gamma$ , which appears to be constitutively bound to that region and is not appreciably enriched upon 4-OHT treatment. Primers specific for a fragment encoding the 3' end of the luciferase coding region (primers LUC2 & LUC3), located nearly 1.2 kbp downstream from the TK promoter, gave no signal in ChIPs for any of these proteins.

Thus, the KPHBD fusion, KAP1, and HP1 $\alpha$  are strongly recruited to chromatin surrounding both the PAX3 binding site and the basal TK promoter elements of the integrated reporter upon 4-OHT treatment. Moreover, no demonstrable signal and/or enrichment was obtained for fragments encompassing the linked SV40-zeocin<sup>R</sup> cassette, or the unlinked CMV promoter or SV40-neomycin<sup>R</sup> cassette. Thus, recruitment of KAP1 and HP1 $\alpha$  is highly specific for a target gene regulated by a DNA bound KRAB repression domain. Moreover, the association of KAP1 and HP1 $\alpha$  to the chromatin, as measured by ChIP, occurs in a highly localized region.

#### EXAMPLE 7: RECRUITMENT OF KAP1 AND HP1 TO AN ENDOGENOUS TARGET GENE

To test whether a similar phenomenon occurs at an endogenous target gene that is regulated by a naturally occurring KRAB transcriptional repressor, NT2 KRAB zinc finger protein targeted repression of the *Col11a2* gene was studied. As discussed below, the NT2-KRAB zinc finger protein stably represses the expression of endogenous *Col11a2* gene in NIH3T3 cells.

The NT2-*Col11a2* system provides a striking example for a KRAB zinc finger protein regulated endogenous target gene. Cartilage-specific expression of  $\alpha 2$  chain of type XI collagen (*Col11a2*) is controlled by regulatory elements located within the -530-bp promoter sequence of mouse *Col11a2*. Interestingly, a 24-bp sequence from -530 to -507 in the *Col11a2* promoter is able to switch the activity of heterologous neurofilament gene (NFL) promoter from neural tissues to cartilage. A yeast one-hybrid screening system has identified specific binding of a KRAB-zinc finger protein

(NT2) to this 24-bp sequence. Biochemical studies revealed that NT2 expression is inversely correlated with *Col11a2* expression and also that it inhibited the *Col11a2* promoter activity through KRAB domain via binding to the 24-bp sequence through the zinc fingers. Thus, the *Col11a2* serves as a valuable endogenous target gene to study the KRAB-KAP repression pathway.

The mouse NT2 KRAB zinc finger protein comprises a leucine-rich SCAN domain, a KRAB box, and nine C-terminal zinc fingers that mediate sequence-specific DNA binding (See Fig. 4A). A western blot analysis using an  $\alpha$ -NT2 antibody (raised in rabbits by injecting the polypeptide representing amino acids 323-345 as the immunogen) was performed as follows: One hundred micrograms of nuclear proteins extracted from NIH3T3 cells were electrophoresed on a 10% SDS-PAGE and the blot was probed with 1:500 dilution of affinity-purified  $\alpha$ -NT2 antibody. The gel (not shown) detected a considerable amount of the ~75 kDa NT2 protein in the NIH3T3 nuclear extract.

Subsequently, NIH3T3 cells were tested for the expression patterns of *NT2*, *Col11a2*, and its neighboring gene *RXR $\beta$*  transcripts by a sensitive quantitative RT-PCR assay such as described above in Example 6. Total RNAs were isolated from NIH3T3 cells and oligo-dT primed first strand cDNAs were made. In order to detect the levels of NT2 mRNA, the following primers were used in the PCR reaction with the first strand cDNAs as template:

NT2-1 primer (5'-GCCAGGCTAGAAGGGAGG-3') SEQ ID NO: 17

corresponding to nucleotides 971-988, and

NT2-2 primer (5'-GGTGTCTGTTGAGGTTGG-3') SEQ ID NO: 18 corresponding to nucleotides 1441-1424 of the mouse NT2 ORF.

Similarly, the *Col11a2* mRNA levels were monitored using COL-1 primer (5'-GGCCTCAGCCTAGCAGATGG-3') SEQ ID NO: 19 corresponding to nucleotides 421-440 and

COL-2 primer (5'-GGCTTATGAAGTCTTGCTGG-3') SEQ ID NO: 20 corresponding to nucleotides 773-754 of the mouse *Col11a2* ORF.

Finally, *RXR- $\beta$*  mRNA levels were determined using RXR-1 primer (5'-GGCTCTGTGCAATCTGCGGG-3') SEQ ID NO: 21 corresponding to nucleotides 569-588 and

**RXR-2 primer (5'-GTCCACAGGCATCTCCTCAGGG-3')** SEQ ID NO: 22  
corresponding to nucleotides 867-846 of the mouse *RXR-β* ORF.

Constant volumes of aliquots removed at cycles 19, 25, 30, 35 or 40 were run on a 1.5% agarose gel. RNA transcripts for the 470 bp *NT2* were detected in cycles 30, 35 and 40. RNA transcripts for the 352 bp *Coll1a2* were not detected. RNA transcripts for the 298 bp *RXRβ* were detected in cycles 35 and 40. Fig. 4B shows the genomic structure of *RXRβ* and *Coll1a2* genes, including the oligonucleotides used in ChIP-PCR experiments, their relative locations, the sizes of the amplified fragments and the exons of *Coll1a2* gene.

Though abundant expression of NT2 transcript was detected, no *Coll1a2* expression was observed. Interestingly, the *RXRβ* gene which is present upstream of *Coll1a2* is highly expressed. This experiment provided evidence that NT2 KRAB zinc finger protein directly and specifically regulates the *Coll1a2* gene expression.

Subsequently, using extensive chromatin immunoprecipitation experiments, the presence of KRAB-KAP1 repression components at the NT2 binding site, promoter sequences, and proximal, distal coding regions of *Coll1a2* gene was analyzed. ChIP analysis of the components of KRAB repression machinery involved preparing soluble, sonicated chromatin prepared from formaldehyde cross-linked NIH3T3 cells and immunoprecipitating the chromatin with the antibodies preimmune (PI) IgG, α-NT2, αRBCC (KAP-1), α-CT (KAP-1), αHP1α, αHP1γ, α-SETDB1, and α-H3MeK. PCRs were carried out on the input and bound, immunoprecipitated DNAs using primers C1 through C8 from Fig. 4B:

**C1 (5'-GGATGCTGCCACGGCCTGAGG-3')** SEQ ID NO: 34 and

**C2 (5'-GGGTCTGCCAGGAGCCTGTGG-3')** SEQ ID NO: 35 primer-pair flank the NT2 binding site (~245 bp), present in the *Coll1a2* promoter-enhancer region.

**C3 (5'-GGGTCGCTATCTATAGCTGG-3')** SEQ ID NO: 36 and

**C4 (5'-GTCCTTTCACACCACGGCAG-3')** SEQ ID NO: 37 primer-pair flank the transcription start site and amplifies the promoter region (~221 bp) of *Coll1a2* gene.

**C5 (5'-GGCCTCAGCCTAGCAGATGG-3')** SEQ ID NO: 19 and

**C6 (5'-GAGCACTCAGACCTTCCAGAGG)** SEQ ID NO: 38 primer-pair amplifies the proximal coding region (~501 bp).

**C7 (5'-GGAACATCAGGTGGTGACGG-3')** SEQ ID NO: 39 and



C8 (5'-GCAGCCCATCCTTCCCTGCAGG-3') SEQ ID NO: 40 primer-pair amplifies the distal coding region of the *Col11a2* gene (~245 bp).

Abundant presence of NT2 was observed at the binding site region. The presence of NT2 was negligent at the promoter region; however KAP1, HP1 $\alpha$ , HP1 $\gamma$ , and SETDB1 proteins were abundantly present. The promoter region was also highly enriched in H3MeK9. Significant hypoacetylation of histone H3 was observed in the same region (data not shown).

These results were strikingly similar to that we observed with engineered KRAB repressor system. Further, these observations provide evidence that a localized heterochromatin structure is generated at the target locus, which may be responsible for stable silencing of *Col11a2* expression.

#### EXAMPLE 8: SPATIAL RE-LOCALIZATION OF THE LUCIFERASE TRANSGENE TO CONDENSED CHROMATIN

Recent reports suggest that gene silencing may be accompanied by spatial re-localization of a gene in the interphase nucleus to a compartment of condensed chromatin (Brown *et al.*, 1997 *Cell*, 91:845-854; Cobb *et al.*, 2000 *Genes Dev.*, 14:2146-2160). Murine cells are ideal for these experiments because the large blocks of pericentromeric heterochromatin form readily identifiable islands in the interphase nucleus.

Fluorescence *in situ* hybridization (FISH) analysis was performed on the KPHBD21 cell line using a biotinylated probe for the CD19-TK-LUC luciferase reporter plasmid. The cells were counter-stained with either DAPI or Hoechst dyes, which preferentially bind to the repetitive A-T-rich sequences that are highly enriched in this pericentromeric heterochromatin structures. For detection of the luciferase gene, either 0.1% ethanol (-OHT) or 500 nM 4-OHT (+OHT) treated KPHBD21 cells were fixed in 4% para-formaldehyde (PFA). When FISH was combined with immunocytochemistry the cells were first immunostained with  $\alpha$ -HP1 $\alpha$  antibody as previously described (Ryan *et al.*, 1999, *Mol. Cell Biol.*, 19:4366-4378), and then re-fixed in 4% PFA to crosslink the bound antibodies. Cells were then permeabilized in 0.2% Triton-X-100 and treated with RNase (Boehringer Mannheim, 100  $\mu$ g/ml in PBS for 30 min at 37°C). After equilibration in 2 $\times$  SSC, cells were dehydrated in an

ethanol series (70%, 80% and 100% ethanol for 3 minutes each at  $-20^{\circ}\text{C}$ ) and air-dried. The hybridization mixture was prepared as follows: The CD19-TK-LUC plasmid was labeled with biotin-16-dUTP by nick translation. The DNase concentration was adjusted to yield probe DNA with a fragment length of 200-500 bases. The probe DNA was dissolved at 10 ng/ $\mu\text{l}$  in 50% formamide in  $2\times$  SSC containing 10% dextran sulfate, 100 ng/ $\mu\text{l}$  salmon sperm DNA (Gibco BRL), 1  $\mu\text{g}/\mu\text{l}$  yeast tRNA (Sigma) and 1.5  $\mu\text{g}/\mu\text{l}$  Cot-1 DNA (Gibco BRL). Both the probe and the cells were simultaneously heated at  $91^{\circ}\text{C}$  for 4 minutes to denature DNA and incubated overnight at  $37^{\circ}\text{C}$ . After hybridization, specimens were serially washed at  $37^{\circ}\text{C}$  with 50% formamide in  $2\times$  SSC ( $2\times$  15 min),  $2\times$  SSC (10 min) and  $0.25\times$  SSC ( $2\times$  5 min). Hybridized probes were detected with FITC-avidin (Vector Laboratories; 1:500 in  $4\times$  SSC plus 0.5% BSA) and the signals were amplified using biotinylated  $\alpha$ -streptavidin (Vector Laboratories, 1:250) followed by another round of FITC-avidin staining. Finally, cells were equilibrated in PBS, stained for DNA with either DAPI (2 ng/ml) or Hoechst (2 ng/ml) and mounted in Fluoromount G (Southern Biotechnology Associates, Inc).

Confocal images of cells were obtained using a Leica confocal laser-scanning microscope. Two channels were recorded simultaneously if no cross talk could be detected. In the case of strong FITC labeling, sequential images were acquired with more restrictive filters to prevent possible breakthrough of the FITC signal into the red channel. Both acquisition modes resulted in the same images. The Leica enhancement software was used in balancing the signal strength and images were scanned 8-fold to separate signal from noise. Alternatively, cells were analyzed with a Leitz Fluovert inverted microscope equipped with a digital camera. Images were obtained using software from QED Imaging (Pittsburgh, PA). Quantitation of luciferase FISH signals, and their association with DAPI or Hoechst-stained spots were done by two independent investigators.

Representative cells were analyzed by FISH. The 4-OHT dependent association of the luciferase transgene was shown with AT-rich condensed chromatin regions. The KPHBD21 cells were seeded onto glass coverslips and treated with either 0.1% ethanol (-OHT) or 500 nM 4-OHT (+OHT) for 48 hrs. Interphase cells

were analyzed by FISH with a luciferase probe and nuclear DNA visualized by counter-staining with DAPI. The results (not shown) were observed as the percent of luciferase FISH signals that were associated (whitish green dots) with the DAPI counter-stained AT-rich sequences in condensed chromatin.

5           The 4-OHT dependent association of the luciferase transgene with AT-rich condensed chromatin regions was also observed. The KPHBD21 cells were seeded onto glass coverslips and treated with either 0.1% ethanol (-OHT) or 500 nM 4-OHT (+OHT) for 48 hrs. Interphase cells were analyzed by FISH with a luciferase probe and nuclear DNA visualized by counter-staining with bis-benzimide (Hoechst 33258)  
10       The results were observed as the percent of luciferase FISH signals that were not associated with the Hoechst counter-stained AT-rich sequences in condensed chromatin.

          A single locus was observed in the KPHBD21 line in every cell examined, strongly suggesting a single integration site for the reporter plasmid (data not shown).  
15       Under growth conditions where the luciferase gene is highly expressed, the FISH signals were most frequently found to be spatially distinct from the condensed chromatin islands as single green dots. However, following addition of 4-OHT and subsequent gene silencing of CD19-TK-LUC, there was a clear association of the majority of luciferase FISH signals with the condensed chromatin territories (data not  
20       shown). This pattern was quantitated by analyzing many cells counter-stained with either DAPI or Hoechst.

          Table II illustrates the number of cells analyzed for each condition of the two FISH experiments described above using an inverted fluorescence microscope. The results are presented as the percent of luciferase FISH signals that were either  
25       associated (whitish green dots) or not associated (green dots) with the DAPI or Hoechst counter-stained AT-rich sequences in condensed chromatin. These results were confirmed in multiple independent experiments.

Table II

Experiment	4-OHT	No. cells analyzed	% of FISH signals	
			Associated with counterstain	Not associated with counterstain
Luciferase-FISH with DAPI counterstain	-	90	27	73
	+	156	69	31
Luciferase-FISH with Hoechst counterstain	-	82	29	71
	+	188	72	28

Further, an immunofluorescent staining of NIH3T3 cells using antibodies raised against KRAB domain stained the DAPI spots indicating a physical association of the endogenous KRAB ZFPs with the condensed chromatin regions and also implicating a role for them in the heterochromatin-mediated gene silencing (data not shown) (Matsuda *et al.*, 2001 4 *J. Biol. Chem.*, 276:14222-14229).

The luciferase FISH experiments were performed after immunostaining the KPHBD21 cells with the HP1 $\alpha$  antibodies (immuno-FISH). These antibodies predominantly stained the heterochromatic islands detected by the DNA dyes, and 4-OHT dependent spatial recruitment of the luciferase FISH signals to the HP1-rich regions was observed (data not shown). These FISH results are consistent with the chromatin compaction and ChIP results reported above. Thus, spatial re-localization of the reporter gene accompanies KRAB-KAP1-HP1 dependent silencing in this system.

#### EXAMPLE 9: TRANSIENT EXPOSURE OF THE LUCIFERASE REPORTER TRANSGENE TO KRAB-PAX3-HBD INDUCES STABLE SILENCING

A hallmark property of HP1-dependent gene silencing, defined by position effect variegation (PEV) in flies, is that the silenced state of the reporter gene is stably maintained through many cell division and differentiation events (Lu *et al.*, 1998 *Dev.*, 125:2223-2234). To determine if there is a stable component to KRAB-KAP1-HP1-mediated gene silencing in the KPHBD21 cell line, 4-OHT wash out experiments were performed as follows:

Cells were seeded at  $0.5 \times 10^4$  in 60-mm dishes. Duplicate dishes of cells were treated with either 500 nM 4-OHT (+OHT) or 0.1% ethanol (-OHT) for the indicated time. At the end of the induction period, one dish from each of the -OHT

and +OHT treated cells was harvested while the remaining dishes were subjected to three washes in a 24-hour period each day. A wash was defined as two changes of 4 ml of DMEM followed by addition of 4 ml of complete growth medium. At the end of 4th day, the cells were harvested by trypsinization and re-plated into new dishes.

5 Cell washings were continued for an additional 4 days. Cells were harvested daily and normalized luciferase activities determined.

The inventors took advantage of the fact that 4-OHT is readily removed from cells and that its effects on HBD fusion proteins are readily reversible (Boehmelt *et al.*, 1992 *EMBO. J.*, 11:4642-4652; Eilers *et al.*, 1989 *Nature*, 340:66-68; Fialka *et al.*, 1996 *J. Cell Biol.*, 132:1115-1132; Pelengaris *et al.*, 1999 *Mol. Cell.*, 3:565-577; 10 Ronchini and Capobianco, 2000 *Oncogene*, 19:3914-3924). These experiments were designed to determine if KRAB-KAP1-HP1-mediated gene silencing could be maintained at the locus in the absence of a DNA binding activity, i.e., the KRAB-PAX3-HBD fusion protein. Replicate dishes of KPHBD21 cells were treated with 15 either 0.1% ethanol or 500 nM 4-OHT for 24 hours (Fig. 6A; filled bars), then either harvested, or extensively washed to remove the hormone. The cells were either harvested, or subjected to an additional wash at twenty-four hour periods thereafter up to 4 days. A twenty-four hour 4-OHT treatment produced about 24-fold reduction in luciferase activity. Twenty-four hours after hormone removal, the luciferase activity 20 was only 2 to 3-fold lower than a duplicate dish, which had not been treated with 4-OHT, thus showing substantial reversal of the repressed state and verifying that the hormone was removed. However, dishes harvested at 2, 3, and 4 days after 4-OHT removal still maintained a substantially lower basal luciferase level than an untreated duplicate dish. The inventors verified that 4-OHT had no effect on the growth rate of 25 4-OHT treated cells and that the cells went through about four rounds of cell division during the 4-day time period. This result was even more pronounced when the initial 4-OHT treatment was for 48 hours (Fig. 6A; patterned bars): dishes maintained in the presence of hormone (and sequentially washed each day) for 4 days showed ~6-fold lower basal luciferase activity compared to untreated dishes. This result provides 30 evidence that gene silencing first initiated by the KPHBD fusion protein may be stably maintained in the absence of the DNA binding component.

Longer-term 4-OHT treatment and washout studies (Fig. 6B) were then performed. A two, four, or six day 4-OHT treatment of KPHBD21 cells strongly repressed the reporter luciferase activity (>50 fold reduction in luciferase activity at 6 days). However, for each treatment, activity was substantially, but not completely recovered compared to untreated duplicate dishes. This effect was still evident at 8 days post-4-OHT removal, a time course, which included extensive washing of each dish everyday, and included a trypsinization and re-plating at semi-confluent cell densities. Similar results were observed with the KPHBD29 and KPHBD36 clones (data not shown). As expected, no repression was observed in two cell clones containing KRAB(DV)-PAX3-HBD fusion protein (Fig. 6C). Moreover, though substantial repression was observed in each of two SNAG-PAX3-HBD cell clones after a 4-day hormone treatment, the luciferase activity was rapidly and fully recovered following 4-OHT removal (Fig. 6D). The inventors also observed full reversal of repression in cell clones expressing either the engrailed, BTB-POZ or WT1 repression domain fusion proteins. Thus, this effect may be specific for the KRAB-KAP-HP1 mechanism of gene silencing.

#### EXAMPLE 10: KRAB-KAP1-HP1 MEDIATED GENE SILENCING IS MITOTICALLY HERITABLE OVER MANY CELL GENERATIONS

Variegated gene expression in *Drosophila*, also observed as sectoring in yeast colonies, reflects the ability of HP1 to maintain the silenced state over many cell generations in a clone of cells. To determine if a similar effect occurred at a euchromatic gene bound by HP1, a clonal analysis of luciferase activities in cells that were transiently pulsed with 4-OHT was performed as follows. The experimental scheme is depicted in Fig. 7A.

Since the original KPHBD21 cell clone had been in culture for an extended period and thus might display clonal heterogeneity, single-cell subclones of KPHBD21 were generated by limiting dilution and grown to mass population. Indeed, among ~100 sub-clones from the KPHBD21 population, a fraction (~10) showed low basal luciferase activity; these might have arisen either due to random epigenetic transgene inactivation that is frequently observed or due to transient

leakiness of the KPHBD protein. The majority of the clones showed high basal luciferase activities similar to the original population.

Duplicate dishes of three of these independent sub-clones (KPHBD21-8, KPHBD21-39 and KPHBD21-49) that express high-level luciferase activities were treated with either 500 nM 4-OHT (+OHT) or 0.1% ethanol (-OHT) for 4 days continuously. Immediately following treatment all dishes were extensively washed, the cells were harvested by trypsinization and subjected to a limiting dilution cloning. Subsequently, wells that contained a single cell were scored, grown to mass population and tested as whole cell lysates for basal (normalized) luciferase activity (Fig. 7B). Single cell progeny were propagated for approximately 40 doublings in normal growth medium.

For sub-line KPHBD21-08, 69 untreated subclones were isolated: their normalized luciferase activities showed a mean of  $\sim 10^5$  light units/O.D. of protein and varied less than 7 fold from lowest to highest. However, sub-line KPHBD21-08 cells that had received a 4-day pulse of 4-OHT prior to single cell cloning yielded a set of sub-clones with a dramatic skewing of activities. More than one-third of 65 clones yielded basal luciferase activities substantially lower than the lowest sub-clone derived from the untreated population. Remarkably, a portion of the clones showed barely detectable luciferase activities. There was no difference in cloning efficiency for treated or untreated KPHBD21-08 cells and the clones that were obtained for each showed similar doubling times (data not shown). A statistical analysis using Fisher's test (F-test), which compares the variances of two samples, yielded a highly significant F value of 0.00000003. Identical results were obtained with sub-clones KPHBD21-39 and KPHBD21-49, where the F values comparing luciferase activity in untreated and treated populations were  $F = 0.00074547$  and  $F = 0.00014801$  respectively.

This experiment was repeated twice with three independent clones and identical results were obtained. The sub-clones maintain resistance to zeocin and G418 and thus have not deleted the transgenes. Moreover, the stably repressed clones have maintained the silent state for approximately four-months in culture (Fig. 7C). These results provide evidence that transient recruitment of the KRAB-KAP1-HP1

complex to a euchromatic gene produces a silenced state that is mitotically heritable and does not require a persistent DNA binding component.

In a similar manner, the silent (clone #39-40) and active (clone #39-45) clones that were selected for detailed characterizations were tested for their basal luciferase activities at different time intervals (approximately at two-month intervals).

#### EXAMPLE 11: MOLECULAR CHARACTERISTICS OF THE TARGET LOCUS IN THE SILENCED AND EXPRESSED CLONES

Since persistence of the silenced and expressed states in clonal cell populations was observed for many generations, the molecular basis for this phenomenon was analyzed. A comprehensive approach consisted of 1) ChIP assays, 2) treatment with chemical inhibitors that are known to affect DNA methylation and histone deacetylase function, 3) transient activation of the silent locus using transcriptional activators, and 4) direct analysis of the DNA methylation status of the reporter genes.

##### A. ChIP Assays

First, the HSV-TK promoter regions of both cl39-45 (expressed) and cl39-40 (silent) clones were analyzed by ChIP assays using a battery of antibodies directed against the molecular components of KRAB-KAP1-HP1 repression pathway. The normalized luciferase activities of the silent and expressed clones determined prior to the ChIP assays were measured (Fig. 8A). Later, the cross-linked chromatin fractions obtained from silent (clone #39-40) and active (clone #39-45) clones were tested in ChIP experiments by using  $\alpha$ -PAX3 IgG, affinity purified  $\alpha$ -KAP1,  $\alpha$ -SETDB1,  $\alpha$ -HP1 $\alpha$ ,  $\alpha$ -HP1 $\gamma$  and  $\alpha$ -H3-MeK9 (Upstate Biotechnology) antibodies. The input and the immunoprecipitated DNAs were amplified using TKP1 and LUC1 primers referred to in Fig. 3 and the samples were tested by Southern blot analysis. The autoradiographic signals (not shown) and the fold difference were obtained by phosphorimager analysis. See Table II.

A strong enrichment of KAP1 (11.66 fold), SETDB1 (10.33 fold), HP1 $\alpha$  (6.93 fold), H3-MeK9 (5.3 fold) and a moderate enrichment of HP1 $\gamma$  (6.93 fold) proteins were observed in the silent clones. However, the level of the KRAB-



PAX3-HBD component at the transgene was very low and similar in the silent and active clones.

**B. Transient Activation of the Silent Locus by a PAX3-VP16 Fusion**

To verify that the PAX3 DNA binding sites are present and accessible  
5 in the silent clone, a VP16-PAX3DBD plasmid that functions as a powerful activator  
was transfected into the silent clone. The VP16 AD-PAX3 DBD (VPDBD)  
expression plasmid was constructed as follows: The VP16 acidic activation domain  
(amino acid residues 1-98) was amplified by PCR from pVP16 plasmid by using  
appropriate primers that introduced a 5' Hind III and a 3' BamH I sites. The PCR  
10 product was cloned at the Hind III and BamH I sites in pcKP plasmid thus replacing  
the KRAB domain, thereby fusing it in frame with the PAX3 DNA binding domain  
(amino acids 99-479). Different concentrations (0, 1, 2, 3, or 4  $\mu$ g) of VPDBD  
plasmid were transiently transfected into the silent clone (clone #39-40) using  
lipofectAMINE. After 24 hrs, the lysates were assayed for luciferase activities and  
15 the normalized light units were determined based on the  $\beta$ -galactosidase values.

Normalized activity with  $\beta$ -galactosidase light units are reported in Fig.  
8B. The observed dose-dependent activation of the silent luciferase locus suggest that  
the molecular components of the KRAB-KAP1-HP1 repression pathway are  
constitutively present at the silenced locus and also that the maintenance of the silent  
20 state is not due to persistent DNA binding of the KRAB-PAX3-HBD protein.

**C. 5-Azacytidine and Trichostatin-A experiments**

Chemical inhibitors such as 5-azacytidine (5AZA) and trichostatin-A  
(TSA) were employed to reactivate the silent locus. Silent (clone #39-40) and active  
(clone #39-45) clones were seeded at  $1 \times 10^5$  cells/60 mm dish in multiple dishes.  
25 Duplicate dishes were treated either with vehicle (0.1% ethanol or PBS) or singly with  
5AZA or TSA or sequentially with both of them or in combination in a sequential  
order for indicated durations as shown in Fig. 8C. The duration of treatment as well  
as the concentrations used are also provided. During sequential treatment, medium  
containing the first compound was removed and replaced with the medium containing  
30 the second compound without any intermittent washes. At the end of chemical  
treatment, the lysates were tested for luciferase activities and protein concentrations  
and normalized light units were determined.

Treatment with either 5AZA or TSA alone did not significantly reactivate the locus. However, a sequential treatment with 5AZA followed by TSA was highly synergistic in reactivating the silent locus (Fig. 8C). The synergy and sequence requirement has been observed before (Cameron, E. E. *et al.*, 1999 *Nat Genet*, 21: 103-107) and strongly suggest that the silent state may be maintained due to DNA methylation.

#### D. DNA Methylation Analysis

The methylation status of the TK promoter regions of the silent and active clones was determined by methylation-specific PCR (MS-PCR) and sodium bisulfite-genomic sequencing.

##### 1. Methylation-specific PCR (MSPCR):

Genomic DNAs were extracted from the silent (clone #39-40) and active (clone #39-45) clones and 1 µg of each was treated with sodium bisulfite to convert the unmethylated cytosines to uracil (Herman, J. G. *et al.*, 1996 *Proc. Natl. Acad. Sci. U S A*, 93: 9821-9826). These DNAs were PCR amplified using unmethylated sense (US): 5' GTTGATTTGGGTATTGAGTTTGAG 3' SEQ ID NO: 41 or methylated sense (MS): 5' GTCGATTCGGGTATCGAGTTTCG 3' SEQ ID NO: 42 primers (present in the TK promoter) and an antisense primer (UMA): 5' GGTTTTATTTTTTAGAGGATAGAATGG 3' SEQ ID NO: 43, present near the start of luciferase cDNA. These MS-PCR primers were designed that target the CpG-rich regions in the TK promoter. One set of primers was designed to only anneal to the bisulfite-modified genomic DNA if the CpG residues were methylated. A similar primer pair was designed that would only anneal if the CpGs lacked methylation. The PCR products were run on an 1.5% agarose gel and photographed (data not shown).

The methylation-specific primer set generated a strong amplification product using DNA from the silent clone. However, no product was seen using DNA from the active clone. Conversely, the unmethylated specific primers only produced a product from the active clone whereas no product was seen in the silent clone. These results suggest that the silent and active clones have different CpG methylation profiles at the promoter transgene.

## 2. Sodium bisulfite-genomic sequencing:

To further map regions of methylation, cloned PCR products derived from bisulfite treated genomic DNAs were sequenced. Genomic DNAs extracted from the silent (clone # 39-40) and active (clone #39-45) clones were treated with sodium bisulfite as above. The sodium bisulfite-modified DNAs were PCR amplified using UMS1 (sense): 5' GTTTTAGTGTTTTATGTTTATAGG 3' SEQ ID NO: 44 and UMA1 (antisense) primers (see Fig. 8D) that were present in regions free of any CpG residues (to avoid any preferential amplification). The PCR products were TA-cloned into pCR II vector (Invitrogen). About 18 recombinant clones of each were sequenced in both directions and the results are presented for seven representative clones in each category in Fig. 8D.

Surprisingly, the active clone showed strong non-random CpG methylation that was highly restricted to the region immediately surrounding the transcription start site. However, the silent clone showed dramatically enhanced density of CpG methylation that was spread both 5' and 3' of the transcription start site. A total of 12 new CpG sites were methylated in the silenced clone. Interestingly, these additional methylated sites were distributed in a highly non-random fashion in the silent TK promoter (Fig. 8D). Together, the sequencing and MS-PCR results provide evidence that DNA methylation contributes to the mitotically heritable gene silencing.

### EXAMPLE 12: KAP-1 COREPRESSOR INTERACTS WITH THE SETDB1 PROTEIN *IN VIVO*

#### A. Plasmids

Full-length human SETDB1 (KIAA0067) was obtained from the Kazasu DNA Research Institute. Coding sequences for SETDB1 were subcloned NotI/BamHI into pCMV2 (Sigma) to create the CMV driven FLAG-tagged SETDB1 mammalian expression vector. The  $\Delta$ KID (amino acids 570 to 1291) expression construct was created by subcloning a HindIII/BamHI fragment into pCMV2. The  $\Delta$ SET (amino acids 1 to 951) expression construct was created by subcloning a NotI/BglII fragment into pCMV2. Amino acid substitutions in SETDB1 (R643V, CC 729, 731 LP, H1224K, C1226A, and C1279Y) were created using Quick Change PCR

mutagenesis strategies (Stratagene). For protein expression in *Escherichia coli*, a 2.2 kbp BamHI fragment encoding amino acid 661 to 1291 of human SETDB1 was subcloned into pGEX-5X-1. Similarly, a 2.6 kbp XhoI/SalI fragment encoding amino acids 585 to 1291 was subcloned into pGEX-5X-1.

5                   B.     Antigen Production

For antigen production, a 1.4 kbp XhoI fragment encoding amino acids 1 to 377 from pACT-KIP41 was subcloned into pGEX-4T-1. Previously described PHD finger and bromodomain mutations in KAP-1 were subcloned into the SmaI site of pBTM116 (Capili *et al.* 2001 *EMBO J.*, 20:165-77; Schultz *et al.* 2001, cited  
10 above). GST-Histone H3 bacterial expression plasmids were previously described (Tachibana *et al.* 2001 *J. Biol. Chem.*, 276: 25309-25317). Bacterial protein expression and *in vitro* GST binding assays were done as previously described (Ryan *et al.* 1999 *Mol. Cell Biol.*, 19:4366-4378; Lechner *et al.* 2000, *Mol. Cell Biol.*, 20:6449-6465). Appropriate reading frame fusions and integrity of flanking sequences  
15 for all constructs created by PCR were confirmed by DNA sequence analysis of both strands.

C.     Yeast Two Hybrid Screen

To identify effectors of KAP-1 directed transcriptional repression, the PHD finger and bromodomain of KAP-1 were used as bait in a yeast two-hybrid  
20 screen. The yeast two-hybrid system as modified by S. Hollenberg was used for all yeast experiments. A human oligo-dT-primed B-cell cDNA library was screened according to the methods of Jensen *et al.*, 1998, *Oncogene*, 16:1097-112.

A novel isoform of the Mi-2 $\alpha$  subunit of the NuRD histone deacetylase complex was previously identified that bound to this bipartite repression domain  
25 (Schultz *et al.* 2001, cited above). This experiment demonstrates that KAP-1 specifically associated with two independent overlapping amino acid sequences (KIP21 and KIP41) that are encoded by the putative histone methyltransferase (HMTase) gene, *SETDB1* (Harte *et al.* 1999 *Cytogenet. Cell Genet.*, 84:83-86). A gel (not shown) illustrated that the KAP-1 PHD finger and bromodomain interact with  
30 Mi-2 $\alpha$  (KIP54) and SETDB1 (KIP21 and KIP 41).

Mutations in the KAP-1 PHD finger and bromodomain that are deleterious to the transcriptional repression activity of the KAP-1 PHD finger and

bromodomain significantly impaired the association between KAP-1 and either SETDB1 or Mi-2 $\alpha$  (Schultz *et al.* 2001 cited above). Some of the mutations differentially affect Mi-2 $\alpha$  and SETDB1 binding which raises the possibility that Mi-2 $\alpha$  and SETDB1 may bind to different surfaces of the KAP-1 protein.

5 *In Vivo* Assay

To confirm the association between KAP-1 and SETDB1 *in vivo*, a full-length FLAG-epitope tagged expression vector and immuno-purified SETDB1 were generated from transfected HEK293 cells as follows. HEK293 cells were transiently transfected with lipofectamine and nuclear extracts were prepared 36-48  
10 hours post-transfection, as previously described (Ryan *et al.* 1999 cited above). Five to ten milligrams of nuclear extract adjusted to 100 mM NaCl were incubated with 100  $\mu$ g of anti-FLAG M2 (Sigma) for 2-4 hours at 4°C. Immune-complexes were washed 3X with BC500 (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 0.2 mM EDTA, 10 % Glycerol, 0.2 mM PMSF, 0.2% Tween 20), 1X with BC100 (20 mM Tris-HCl, pH  
15 8.0, 100 mM NaCl, 0.2 mM EDTA, 10 % Glycerol, 0.2 mM PMSF, 0.2% Tween 20), and eluted with 400 mg/ml FLAG M2 peptide. Eluted proteins were resolved on a 4-12% NuPAGE gel in MOPS running buffer (Invitrogen). Proteins were visualized by silver staining or Western blotting to PVDF as previously described (Ryan *et al.* 1999 cited above).

20 For endogenous SETDB1 immunoprecipitation studies, 100  $\mu$ g of a DEAE bound, 0.1 M phosphocellulose elution of HeLa nuclear extract was incubated with 5  $\mu$ g of affinity purified SETDB1 antibody and 5  $\mu$ l of protein G-Sepharose (Pharmacia) for 2 hours at 4°C. Bound immune complexes were washed 3 times with BC100 and twice with HMTase buffer prior to assaying for HMTase activity.

25 The spectrum of polypeptides was subjected to MS/MS peptide analyses using microcapillary reverse phase HPLC nano-spray tandem mass spectrometry on a Finnigan LCQ quadrupole ion trap mass spectrometer.

KAP-1 was definitively identified as a non-stoichiometric, associated polypeptide. A Coomassie blue stain of anti-FLAG immuno-purified SETDB1 from  
30 transfected HEK293 cells analyzed with MS/MS peptide identification, definitively identified overlapping peptides of KAP-1, spanning a significant portion of the KAP-1 open reading frame (amino acids 239 to 790). These peptides were:

	DHQQYQFLEDAVR	SEQ ID NO: 23
	FASWALESDNNTALLLSK	SEQ ID NO: 24
	LIYFQLHR	SEQ ID NO: 25
	FQWDLNAWTK	SEQ ID NO: 26
5	IVAERPGTNSTGPAPMAPPP	SEQ ID NO: 27
	VFPGSTTEDYNLIVIER	SEQ ID NO: 28
	EEDGSLSLDGADSTGVVAK	SEQ ID NO: 29
	LSPPYSSPQEFAQDVGR	SEQ ID NO: 30
	ADVQSIIGLQR	SEQ ID NO: 31

10                   An anti-KAP-1 Western blot of FLAG immunoprecipitates from HEK293 transfected nuclear extracts was generated (data not shown). These results showed that the KAP-1 corepressor interacts with the SETDB1 protein *in vivo*.

15                   The primary amino acid sequence of SETDB1 revealed several interesting signature motifs including a CpG-DNA methyl binding domain of the MeCP2 family, and homology to the SET (SuVar3-9, Enhancer of Zeste, Trithorax) domain (Fig. 9). Interestingly, the SET domain homology of SETDB1 is interrupted by a 347 amino acid insertion to create a bifurcated domain (Harte *et al.* 1999, cited above). This unique insertion is evolutionarily conserved from the human protein to lower eukaryotes, including *C. elegans* and *D. melanogaster*, suggesting that the SET domain may possess functionally separable domains. It has been previously demonstrated that the SET domain homology of SUV39H1 and the two adjacent cysteine rich, (pre-SET and post-SET) regions possess intrinsic histone methyltransferase (HMTase) activity that is dependent on the integrity of all three domains (Rea *et al.* 2000 *Nature*, 406:593-599).

25

**EXAMPLE 13: IN VITRO HISTONE METHYLTRANSFERASE REACTION****A. Protocol**

The histone methyltransferase assays were conducted as follows: In a 40  $\mu$ l reaction volume, enzyme, 5  $\mu$ g of core histones (Roche Biochemicals), 2  $\mu$ g of chicken erythrocyte mononucleosomes, or 5  $\mu$ g of GST-H3N, and 500 nCi of S-adenosyl-( $^3$ H-methyl)-L-methionine ( $^3$ H-AdoMet; 72 Ci/mmol; NEN Life Science Products) were incubated for 1 hour at 37°C in 50 mM Tris, pH 8.5, 20 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM  $\beta$ -mercaptoethanol, and 250 mM sucrose. Reactions were terminated by the addition of 5X SDS-buffer. Histones were resolved on 4-12% NuPage gels in MES running buffer and visualized by Coomassie Blue R250 stain.  $^3$ H-methyl labeling was detected by flourography in 22% PPO solution. Dried gels were exposed to Kodak MRX film. Western Blotting was done as previously described (Ryan et al. 1999, cited above).

**B. SETDB1 proteins**

To test whether SETDB1 possessed intrinsic HMTase activity, two different recombinant GST-SETDB1 fusion proteins that encode the entire putative catalytic domain (amino acids 585 to 1291 and 661 to 1291, respectively) were expressed and purified. To evaluate the role of the pre-SET, SET, and post-SET domains in the HMTase activity, several mutants of SETDB1 were engineered. The HMTase activity of a protein containing a deletion of the post-SET domain and part of the SET homology was significantly impaired, at least in this assay system. Furthermore, single amino acid substitutions at highly conserved residues in each of these subdomains reduced the methylase activity to undetectable levels. However, deletion of the putative KAP-1 interaction domain (KID) domain increased activity. A point mutation in the MBD homology had no effect on this enzymatic activity (data not shown). The deletion of the post-SET and part of the SET homologies and single amino acid mutations at highly conserved residues within the catalytic domain (pre-SET, SET and post-SET) impairs the H3-methylase activity of SETDB1. The anti-FLAG Western blot (not shown) confirmed the expression and FLAG immunopurification of the indicated proteins. The  $\Delta$ KID and  $\Delta$ SET proteins correspond to amino acids 570 to 1291 and 1 to 951 of SETDB1, respectively.

Thus, similar to other members of the SET domain family of histone methyltransferases, SETDB1 requires the pre-SET, SET, and post-SET homologies for full enzymatic activity *in vitro*. Remarkably, the unique 347 amino acid insertion in the SET domain appears to have no effect on the catalytic activity of SETDB1.

These data support previous hypotheses that the SET domain structurally may be composed of two separable functional domains (Katsani *et al.* 2001 *Genes Dev.*, 15:2197-2202). Furthermore, binding to KAP-1 is apparently not required for the methyltransferase activity. Although the contributions of other endogenous polypeptides present in the immuno-purified preparations of SETDB1 cannot be ruled out, purification of SETDB1 to near homogeneity from Sf9 baculovirus infected cell extracts indicates that it is sufficient to mediate histone methylation (data not shown).

#### C. Specificity

To define the site specificity of H3 methylation by SETDB1, a series of purified, recombinant GST-histone tail proteins with several lysine to arginine substitutions were used as substrates (Fig. 10A) (Tachibana *et al.* 2001 cited above). SETDB1 selectively methylates lysine 9 of histone H3. The amino acid sequence of the NH<sub>2</sub>-terminal tail of histone H3 (1-30) is shown with the K4, K9, and K27 residues highlighted as follows.

ARTK<sub>4</sub>QTARK<sub>9</sub>STGGKAPRKQLATKARK<sub>27</sub>SAP SEQ ID NO: 32

Various lysine to arginine mutations in K4, K9, and K27 derived to determine the substrate specificity of SETDB1. For NT, lysine (K) to arginine (R) mutations are present at K4, K9, and K27. For N4, lysine (K) to arginine (R) mutations are present at K9, and K27. For N9, lysine (K) to arginine (R) mutations are present at K4 and K27. For N27, lysine (K) to arginine (R) mutations are present at K4 and K9. For K4R, a lysine (K) to arginine (R) mutation occurs at K4. For K9R, lysine (K) to arginine (R) mutation occurs at K9. For K27R, lysine (K) to arginine (R) mutation occurs at K27.

H3 methylation by SETDB1 was found to be is highly selective for lysine 9. SETDB1 methyltransferase activity is highly specific for K9. Five µg of the corresponding GST-H3N protein was used as substrate in the *in vitro* HMTase assay with FLAG purified SETDB1. Coomassie blue stain showed (not shown) the purified



GST-Histone H3 substrates. Autoradiograph showed (data not shown) corresponding <sup>3</sup>H-methyl-labeled products. Western blot (not shown) confirmed the presence of FLAG-SETDB1 in the HMTase reaction. A substrate (NT) in which K4, K9, and K27 were each mutated to arginine failed to be methylated. Substrates with double lysine to arginine mutations (N4, N9, N27) revealed methylation of a substrate with only K9 (N9) preserved. A substrate with a single arginine substitution at K9 (K9R) confirmed the specificity of SETDB1 for K9.

These data confirm that additional post-translational modifications (i.e. acetylation, phosphorylation, methylation) of the substrate are not required for H3-K9 methylation by SETDB1. Moreover, when K9 was mutated to arginine (K9R), SETDB1 did not change its specificity to K27, despite the fact that this residue lies in a strikingly similar amino acid sequence (TKxxARKS; SEQ ID NO: 33) as K9.

#### D. Assays

Since HP1 proteins bind methylated K9 histone peptides, an experiment was conducted to determine whether SETDB1 could stimulate HP1 binding to the NH2-terminal tails of histone H3. *In vitro* GST binding assays between HP1 $\alpha$  and GST-H3N are performed as follows: GST-H3N substrates were pre-methylated with Flag-purified SETDB1 and 15  $\mu$ M S-adenosyl-L-methionine (Sigma). <sup>35</sup>S-L-methionine labeled *in vitro* translated HP1 $\alpha$  proteins were incubated with the methylated GST-H3N proteins. HP1: histone complexes were eluted by denaturation, resolved on 10% SDS-PAGE gels, and bound HP1 $\alpha$  was visualized by fluorography. Coomassie blue stain (not shown) illustrated the purified, methylated GST-Histone H3 substrates. Methylation of the GST-H3 and GST-N9 substrates by SETDB1 significantly enhanced the efficiency of HP1 $\alpha$  binding to the NH2-terminal tail of histone H3. This binding activity was abolished by a mutation in the chromodomain (V21M) of HP1 $\alpha$  (See Fig. 10A). Furthermore, a mutation in the chromoshadow domain (I165K) that effects the dimerization of HP1 $\alpha$  significantly impaired the HP1: histone interaction (Lechner *et al.* 2000, cited above) (See Fig. 10A). From this series of data it is concluded that SETDB1 is a highly selective histone H3-K9 methylase fully capable of stimulating the binding of HP1 proteins to histone H3.

### E. Effect of Modifications

In order to determine whether pre-existing post-translational modifications (i.e. acetylation, methylation, and phosphorylation) of histone H3 affect the SETDB1 methylase activity, the activity of SETDB1 was tested against a panel of peptide substrates possessing either an individual or a combination of modifications (Fig. 10B). FLAG purified SETDB1 robustly methylated the unmodified H3 substrate, but not a H4 peptide. Interestingly, a peptide substrate methylated at K4 had no apparent effect on this activity. As expected, any modification (methylation or acetylation) of H3-K9 inhibited SETDB1-mediated methylation. Furthermore, phosphorylation of S10 or acetylation of K14 also dramatically inhibited the methylation of the substrate. These observations are similar to that previously observed for the related K9-specific histone H3 methyltransferase SUV39H1, indicating that these proteins likely recognize the substrate in a similar fashion and possess a similar catalytic mechanism. Thus, the ability of SETDB1 to methylate histones within a target locus *in vivo* likely requires coordination with deacetylase complexes and putative histone phosphatases.

Unlike recombinant PRMT1, SUV39H1 and G9a proteins, the recombinant SETDB1 proteins failed to demonstrate any appreciable methylation of core histones (data not shown). However, SETDB1 that was immunopurified from transiently transfected HEK293 cells showed a robust histone H3-specific methyltransferase activity for core histones and mononucleosome substrates. Identical enzymatic activity was observed for a SETDB1 protein expressed and purified to homogeneity from baculovirus infected Sf9 cell extracts (data not shown). Thus, SETDB1 may require post-translational modification or a cellular cofactor(s) in order to function as a histone methylase.

### F. Methylase Activity

To confirm that endogenous SETDB1 possessed methylase activity both monoclonal and polyclonal antibodies were produced that specifically recognize the protein (Fig. 9). Western blot analysis of phosphocellulose-fractionated soluble HeLa nuclear extract revealed that SETDB1 primarily elutes in the 0.1 M and 0.3 M KCl elutions, whereas SUV39H1 is present in the 0.5 M and 1.0 M KCl elutions.

The fractions containing either protein displayed robust H3 methylase activity. Endogenous SETDB1 represents a major histone H3 specific methyltransferase as is shown by the biochemical fractionation of H3 specific methyltransferases from HeLa nuclear extract. HeLa nuclear extract was fractionated  
5 by phosphocellulose (P11) chromatography as previously described (Bochar et al. 2000). HMTase activity was monitored by the *in vitro* methylation assay. Elution of SETDB1 and SUV39H1 from the P11 column was monitored by Western blot analysis (data not shown).

Antibodies against SETDB1 efficiently immunodepleted nearly all the  
10 histone H3 methylase activity from the 0.1 M P11 extract without affecting the H4 activity. 150 µg of the 0.1 M P11 fractionated nuclear extract was incubated with protein A-agarose and either affinity purified anti-GST or anti-SETDB1 IgG. Supernatants and pellets from these immunoprecipitates were assayed for HMTase activity. Coomassie blue stain (not shown) revealed equal amounts of core histone  
15 substrate in each reaction. Supernatants of SETDB1 immunodepleted nuclear extract (not shown) were devoid of H3 HMTase activity. Autoradiographs (not shown) were made of corresponding <sup>3</sup>H-methyl-labeled products. Anti-SETDB1 western blot (not shown) demonstrated efficient immunodepletion of SETDB1 from the 0.1 M P11 extract.

20 Moreover, the pellet of the SETDB1 immunoprecipitate retained a strong histone H3 activity that is comparable to that of FLAG-purified SETDB1 (data not shown).

The peptide eluate of anti-FLAG immunopurified SETDB1 was obtained from transiently transfected HEK293 cells. Histone H3-specific  
25 methyltransferase activity is revealed in an *in vitro* HMTase assay with either core histones or chicken erythrocyte mononucleosomes as substrates. Coomassie blue stain (not shown) showed the loading of histones whose identities are labeled respectively. An autoradiograph demonstrated (not shown) corresponding <sup>3</sup>H-methyl-labeled products. The above experiments strongly suggest that endogenous SETDB1  
30 represents an abundant histone H3 methyltransferase.

**EXAMPLE 14: LOCALIZATION OF SETDB1**

Indirect immunofluorescence staining of asynchronous populations of NIH/3T3 cells was conducted as follows: NIH/3T3 cells were grown on glass coverslips in DMEM medium containing 10% calf serum and immunostained as previously described (Maul *et al.* 1998 *Cell Growth Differ.*, 9:743-755). The murine SETDB1 protein was visualized by indirect immunofluorescence with an antigen-purified rabbit polyclonal antibody diluted 1:200. DNA was counter stained with Hoechst 33258 (Sigma) and coverslips were mounted with Fluoromount G (Fisher Scientific). Cells were visualized with an inverted light microscope (Leica Inc).

This experiment revealed that SETDB1 is localized predominantly in euchromatic regions of interphase nuclei and excluded from nucleoli and islands of condensed chromatin, as determined by Hoechst stain and immuno-staining with a monoclonal antibody to HP1 $\alpha$ . Affinity purified polyclonal SETDB1-specific IgG globally stained euchromatic nuclear territories of interphase nuclei (FITC) of NIH/3T3 cells with little overlap in A-T rich condensed chromatin domains were visualized by Hoechst stain and monoclonal HP1 $\alpha$  IgG (Texas Red).

However, there is significant overlap between SETDB1 and HP1 $\alpha$  in euchromatic regions of the nucleus. Thus, SETDB1 functions independently of SUV39H1/H2 and is one cellular HMTase responsible for global euchromatic H3-K9 methylation maintained in the *Suv39h* double knockout mouse (Peters *et al.* 2001 *Cell*, 107: 323-327). These results suggest that SETDB1 functions to methylate histone H3-K9 in euchromatic territories of the nucleus to facilitate HP1 deposition.

**EXAMPLE 15: CHROMATIN IMMUNOPRECIPITATION (CHIP)****EXPERIMENTS**

To test whether the KRAB-ZFP: KAP-1 repression complex could target H3-K9 methylation of endogenous gene promoters by SETDB1, facilitating the deposition of HP1 proteins to silence gene expression, chromatin immunoprecipitation (ChIP) experiments were done with a cell line that contains a stably integrated, euchromatic luciferase transgene that is subject to KRAB-mediated repression (Fig. 11). This two-plasmid system is based upon the hormone-regulatable DNA binding KRAB domain fusion (Ayyanathan *et al.* 2000, cited above) and a TK-

luciferase reporter transgene as its target. The fusion protein is fully capable of forming a ternary complex with KAP-1 and HP1 (data not shown).

The above described model for HP1-dependent silencing and variegation of a euchromatic gene expression in a mammalian cell line was utilized as described  
5 below to evaluate the role of SETDB1 and histone H3 MeK9 in the stable silencing of the luciferase transgene. Clonal cell lines were used that demonstrated either robust expression of luciferase (cl-49) or nearly complete silencing of the luciferase transgene (cl-74). Two single cell sub-clones containing the heterologous KRAB-PAX3-HBD transcriptional repressor and the integrated luciferase transgene, were  
10 either expressed (cl-49) or stably silenced (cl-74) following hormone treatment. Luciferase activities were measured in subconfluent populations of cells and reported as relative light units/mg of protein.

Chromatin immunoprecipitation (ChIP) experiments were done essentially as previously described with some modifications (Orlando *et al.* 1997 *Methods*, 11:205-  
15 214). Cells were cross-linked with 1% formaldehyde for 20 minutes at 37°C. The cross-linking reaction was quenched by washing the cells several times with cold TBS (50 mM Tris, pH 8.0, 200 mM NaCl). Cells were scraped into cold TBS supplemented with 5 mM butyric acid. Chromatin was enriched for by washing the cells once in 20 mM Tris HCl pH 8, 0.25% Triton X-100, 200 mM NaCl, 10 mM  
20 EDTA, 0.5 mM EGTA, 1 µg/mL aprotinin, leupeptin, pepstatin, 1 mM Benzamidine, 50 mM NaF, 10 mM NaOV<sub>3</sub>, and 5 mM butyric acid. The cells were centrifuged, resuspended in IP buffer (20 mM NaCl, 0.05% DOC, 0.5% Triton X100, 0.5% NP-40, 200 mM NaCl, PIs and 5 mM butyrate) and sonicated to average fragment size of 300-500 bp. Solubilized chromatin was clarified by centrifugation at 12,000Xg and the supernatant was pre-incubated for 2 hours with protein A agarose beads blocked  
25 with Salmon sperm DNA and BSA. Pre-cleared chromatin was incubated with 5 to 10 µg of anti-KAP-1 (Schultz *et al.* 2001, cited above) anti-SETDB1, anti-H3 MeK9 (Upstate Biotechnology) and anti-HP1α (Schultz, unpublished data) for 12-16 hours at 4°C. Immune complexes were bound to protein A agarose beads for an additional 2-3  
30 hours at 4°C. The beads were washed 4 times with IP buffer, two times with high salt buffer (IP buffer with 0.4 M NaCl), once with LiCl buffer (10 mM Tris pH8, 250 mM LiCl, 0.5% NP40, 1% Triton X-100, 0.1% DOC, 5mM EDTA, PIs), and two times

with TE. DNA: protein complexes were eluted from the protein A beads with 50 mM Tris, pH 8.0, 200 mM NaCl, 5 mM EDTA, 1% SDS for 1 hour at RT. The supernatant was transferred to a fresh tube and cross-links reversed at 65°C for 6-12 hours. Samples were treated with 30 µg Proteinase K (Roche Biochemicals) for 2 hours at 55°C, extracted once with phenol, and the DNA precipitated with 2.5 volumes of ethanol plus 20 µg of glycogen as carrier. Precipitated DNA was pelleted, washed once with 70% ethanol, dried, and resuspended in 25 µl of water. DNA was analyzed by PCR using specific primer pairs to promoter sequences of the integrated plasmids.

Gels (not shown) from ChIP experiments demonstrated the colocalization of KAP-1 and SETDB1 at the TK promoter region of the luciferase transgene in the cells where transcription of the luciferase gene has been stably silenced (cl-74). Formaldehyde crosslinked chromatin from cl-49 and cl-74 cells was immunoprecipitated with either affinity purified KAP-1 or SETDB1 IgG. An equal amount of promoter sequence in cl-49 and cl-74 nucleosomal preparations was determined by PCR from 1% of the input chromatin.

In cells containing a silenced luciferase transgene (cl-74), the ChIP experiments indicated that both KAP-1 and SETDB1 readily cross-linked to the luciferase transgene and were significantly co-localized around the TK promoter region of the integrated reporter. In contrast, KAP-1 and SETDB1 were undetectable by CHIP analysis at the TK promoter region in cells (cl-49) displaying significant luciferase activity. Furthermore, little binding of these proteins was observed to the promoter region of the linked Zeocin resistance locus that is nearly 3.0 kbp downstream of the TK promoter, or at the unlinked Neomycin resistance gene present in the same cells. Clonal populations of cells containing only the integrated luciferase transgene failed to demonstrate any localization of KAP-1 or SETDB1 to the TK promoter region, suggesting that a DNA bound KRAB repression module was required for KAP-1 and SETDB1 recruitment (data not shown).

It was next determined whether the localization of SETDB1 to the TK promoter region enhanced H3-K9 methylation and the recruitment of HP1. Gels (not shown) of ChIPs of crosslinked chromatin with KAP-1, SETDB1, HP1α, and MeK9 antiserum illustrated a comparison between these two cell lines and revealed that

HP1 $\alpha$  and its chromatin ligand, H3-MeK9, are enriched in chromatin containing the TK promoter sequences in cells with a hormone-induced, stably silenced transgene. These data suggest that the KAP-1 corepressor functions as a molecular platform that coordinates the sequential recruitment of histone methyltransferases and the  
5 deposition of HP1 at a euchromatic locus to stably silence gene expression that is mitotically heritable.

All publications cited in this specification are incorporated herein by reference.

**CLAIMS:**

1. A method for producing a cloned cell containing a stably silenced target gene, said method comprising the steps of:
  - (a) introducing into a parent cell a nucleic acid molecule expressing a chimeric repressor fusion protein, said repressor protein comprising:
    - i. a first amino acid sequence comprising a Krüppel-Associated Box (KRAB) domain, or a variant thereof which binds the protein KAP1 and has DNA-dependent repressor activity, fused to
    - ii. a second amino acid targeting sequence that binds to said target gene, fused to
    - iii. a switch component, that, in the presence of a ligand or inducer, permits said second amino acid sequence to bind to said target gene; said fusion protein being under the control of regulatory sequences capable of directing expression thereof in said parent cell;
  - (b) culturing said parent cell in the presence of an effective amount of said ligand for a sufficient time to repress expression of said target gene;
  - (c) culturing said ligand-treated cells (b) in the absence of ligand to obtain progeny cells thereof; and
  - (d) isolating from said progeny cells, those cells that do not express said target gene after at least about 25 to about 50 cell population doublings, wherein said isolated, non-expressing cells (d) contain said target gene stably silenced.
2. The method according to claim 1, wherein step (c) comprises single cell cloning said ligand-treated cells (b) in the absence of ligand to obtain single subclones thereof; and step (d) comprises isolating from said subclones, those cells that do not express said target gene after at least about 25 to about 50 cell population doublings.



3. The method according to claim 1, wherein said parent cell is a eukaryotic cell.
4. The method according to claim 3, wherein said cell is a mammalian cell.
5. The method according to claim 4, wherein said mammalian cell is selected from the group consisting of a multipotent stem cell, a pluripotent stem cell, and an embryonic stem cell.
6. The method according to claim 1, wherein said targeting sequence is heterologous to said first amino acid sequence.
7. The method according to claim 1, wherein said targeting sequence is a DNA binding domain that binds to a DNA binding site in said target gene.
8. The method according to claim 7, wherein said DNA binding site is a monomeric, extended non-degenerate DNA sequence in said target gene.
9. The method according to claim 1, wherein said targeting sequence is a protein association motif.
10. The method according to claim 1, wherein said nucleic acid molecule is a recombinant viral vector.
11. The method according to claim 1, wherein said nucleic acid molecule is a non-viral vector.
12. The method according to claim 1, wherein said nucleic acid molecule is naked DNA.

13. The method according to claim 1, wherein said regulatory sequences comprise a promoter selected from the group consisting of a constitutive promoter, an inducible promoter, and a tissue-specific promoter.

14. The method according to claim 1, wherein said switch component is a ligand dependent binding domain obtained from a member of the nuclear hormone receptor superfamily.

15. The method according to claim 14, wherein said receptor is the estrogen receptor or the progesterone receptor.

16. The method according to claim 14, wherein said ligand is a synthetic hormone.

17. The method according to claim 16, wherein said ligand is 4-hydroxytamoxifen (4-OHT).

18. The method according to claim 1, wherein said target gene is a gene encoding a reporter exogenous to said cell.

19. The method according to claim 1, wherein said target gene is a gene endogenous to said cell.

20. The method according to claim 19, wherein said target gene is associated with a disease.

21. The method according to claim 1, wherein said effective amount of ligand is between about 100 to about 1000 nM ligand.

22. The method according to claim 21, wherein said effective amount of ligand is about 500 nM ligand.

23. The method according to claim 1, wherein said sufficient time is greater than 12 hours.
24. The method according to claim 23, wherein said sufficient time is about 24 hours to about 150 hours.
25. The method according to claim 14, wherein said nuclear receptor is an estrogen receptor, said ligand is 4-OHT, said effective amount is about 500 mM ligand and said sufficient time is about 96 hours.
26. The method according to claim 1, wherein said population doublings are 40.
27. The method according to claim 1, further comprising administering multiple different said nucleic acid molecules, each having a different ligand binding domain and a different ligand for silencing of multiple target genes.
28. A cell containing a stably silenced target gene produced by the method of any of claims 1 to 27
29. A method for identifying a test molecule that activates the expression of a stably silenced target gene, said method comprising the steps of:
- (a) contacting a cell of claim 28 or progeny thereof with a test molecule; and
  - (b) monitoring said cells or progeny for expression of at least one said silenced target gene,
- wherein expression of said target gene in the presence of said test molecule indicates that said test molecule reactivates expression of at least one said silenced gene.
30. The method according to claim 29, wherein said at least one silenced target gene is a reporter gene.

31. The method according to claim 29, wherein said test molecule is a molecule that disrupts DNA methylation or histone acetylation at said silenced target gene.

32. A method for manipulating expression of a target gene in a cell, said method comprising the steps of:

(a) introducing to said cell containing said target gene an effective amount of a composition comprising a nucleic acid molecule expressing a chimeric repressor fusion protein, said repressor protein comprising:

i. a first amino acid sequence comprising a Krüppel-Associated Box (KRAB) domain, or a variant thereof which binds the protein KAP1 and has DNA-dependent repressor activity, fused to

ii. a second amino acid targeting sequence that binds to said target gene, fused to

iii. a switch component, that, in the presence of a ligand or inducer, permits said second amino acid sequence to bind to said target gene;

said fusion protein being under the control of regulatory sequences capable of directing expression thereof in said parent cell; and

(b) stably silencing expression of said target gene by transiently contacting said cell (a) with an effective amount of said ligand for a sufficient time to stably suppress expression of said target gene in subsequent progeny of said cell in the absence of said ligand.

33. The method according to claim 32, wherein said ligand causes the formation at the target gene of a complex comprising the KAP1 protein, the HP1 protein and the SETDB1 enzyme.

34. The method according to claim 32, further comprising the step of:

(c) reactivating expression of said silenced target gene in said progeny cells by contacting said cells with a molecule that disrupts the binding at the target gene of said complex.

35. The method according to claim 32, wherein said target gene is exogenous to said cell.

36. The method according to claim 32, wherein said target gene is endogenous to said cell.

37. The method according to claim 36, wherein expression of said target gene results in a disease or disorder.

38. The method according to claim 32, wherein said introducing step (a) occurs *ex vivo*.

39. The method according to claim 38, wherein said delivering step (b) occurs *ex vivo*.

40. The method according to claim 32, wherein said introducing step (a) comprises administering said composition to said cell in a mammalian patient by a suitable route of administration.

41. The method according to claim 40, wherein said delivering step (b) comprises administering said composition to said cell in a mammalian patient by a suitable route of administration.

42. The method according to claim 32, wherein said cell is a stem cell.

43. A method of producing a knock-out, non-human animal in which a selected target gene is stably silenced, comprising the steps of:
- (a) introducing to said cells of said animal containing said target gene an effective amount of a composition comprising a nucleic acid molecule expressing a chimeric repressor fusion protein, said repressor protein comprising:
    - i. a first amino acid sequence comprising the Krüppel-Associated Box (KRAB) domain or a variant thereof which binds KAP1 and has DNA-dependent repressor activity, fused to
    - ii. a second amino acid targeting sequence that binds to said target gene, fused to
    - iii. a switch component, that, in the presence of a ligand or inducer, permits said second amino acid sequence to bind to said target gene; said fusion protein being under the control of regulatory sequences capable of directing expression thereof in said parent cell;
  - (b) transiently administering to said animal an effective amount of said ligand to repress expression of said target gene, wherein expression of said target gene is silenced in subsequent progeny of said cells in said animal in the absence of said ligand.

44. A nucleic acid molecule encoding a chimeric repressor fusion protein, said repressor protein comprising:
- i. a first amino acid sequence comprising the Krüppel-Associated Box (KRAB) domain, or a variant thereof which binds KAP1 and has DNA-dependent repressor activity, fused to
  - ii. a second amino acid targeting sequence that binds to said target gene, fused to
  - iii. a switch component, that, in the presence of a ligand or inducer, permits said second amino acid sequence to bind to said target gene; said fusion protein being under the control of regulatory sequences capable of directing expression thereof in said parent cell.